NRA
WE DO OUR PART
APR 1

MARCH 1935

Journal of Educational Research

Registered June 4, 1929

CONTENTS

		PAGE
A Symposium on the Effects of Measurement on Instruction		481
The Effects of Measurement on Instruction	*A. S. Barr*	481
The Effects of Measurement on Instruction	*W. W. Theisen*	484
Tests: Aims, Procedure, and Types	*Guy M. Wilson*	488
Measurement and Instruction	*L. W. Keeler*	493
Effect of Measurement on Instruction	*Walter S. Monroe*	496
The Effect of Measurement on Instruction	*A. C. Krey*	498
The Normal Relationship Between Measurement and Instruction	*M. R. Trabue*	501
The Effect of Measurement on Instruction	*John M. Stalnaker* and *M. W. Richardson*	504
The Effect of Measurement upon Instruction	*Harl R. Douglass*	508
Coöperative Achievement Testing	*E. F. Lindquist*	511
Summary and Reactions	*Clifford Woody*	520
Editorials: Measurement and Evaluation	*Clifford Woody*	528
Editorials: Educational Diagnosis	*A. S. Barr*	529
Reviews		532
Research Abstracts and Bibliographies		539
Research News and Communications		547

THE JOURNAL OF EDUCATIONAL RESEARCH is indexed in the Education Index

Copyright 1935, by Public School Publishing Co.

Entered as second-class matter at the Post Office at Bloomington, Illinois, under the act of March 3, 1879

JOURNAL OF EDUCATIONAL RESEARCH

EDITORIAL BOARD

A. S. BARR, *Chairman, Professor of Education, University of Wisconsin, Madison, Wisconsin*

E. J. ASHBAUGH, *Dean, College of Education, Miami University, Oxford, Ohio.* Administration, Supervision, and Teacher Training.

HARRY J. BAKER, *Director, Psychological Clinic, Detroit Public Schools, Detroit, Michigan.* Child Welfare, Guidance, and Development.

LEO J. BRUECKNER, *Professor of Education, University of Minnesota, Minneapolis, Minnesota.* Elementary Education.

CARTER V. GOOD, *Professor of Education, University of Cincinnati, Cincinnati, Ohio.* Teaching, Learning, and Individual Differences.

KAI JENSEN, *Assistant Professor of Education, University of Wisconsin, Madison, Wisconsin.* Preschool Education.

CHARLES W. KNUDSEN, *Professor of Education, George Peabody College, Nashville, Tennessee.* Secondary Education.

A. K. LOOMIS, *Principal, University High School, University of Chicago, Chicago, Illinois. Curriculum Construction.*

CLIFFORD WOODY, *Director, Bureau of Educational References and Research and Professor of Education, University of Michigan, Ann Arbor, Michigan.* Measurement, Statistics, and Research.

To save time, manuscripts may be sent directly to members of the editorial board as indicated above.

ASSOCIATE EDITORS

FRED C. AYER, *Professor of Educational Administration, University of Texas, Austin, Texas*

FOWLER D. BROOKS, *Head of Departments of Education and Psychology, DePauw University, Greencastle, Indiana*

WILLIAM A. BROWNELL, *Professor of Educational Psychology, Duke University, Durham, North Carolina*

B. R. BUCKINGHAM, *Lecturer in Education, Harvard University, Cambridge, Massachusetts*

W. W. CHARTERS, *Director, Bureau of Educational Research and Professor of Education, Ohio State University, Columbus, Ohio.*

VIRGIL E. DICKSON, *Assistant Superintendent of Schools and Director of Bureau of Research and Guidance, Berkeley Public Schools, Berkeley, California.*

EDGAR A. DOLL, *Director, Department of Research, Training School, Vineland, New Jersey*

WILLIAM S. GRAY, *Director, Teacher Training, University of Chicago, Chicago, Illinois*

CHARLES E. GREENE, *Assistant Superintendent of Schools, Denver Public Schools, Denver, Colorado*

M. E. HAGGERTY, *Dean, College of Education and Professor of Education, University of Minnesota, Minneapolis, Minnesota*

V. A. C. HENMON, *Chairman, Department of Psychology and Professor of Psychology, University of Wisconsin, Madison, Wisconsin*

KARL J. HOLZINGER, *Professor of Education, University of Chicago, Chicago, Illinois*

ERNEST HORN, *Director, University Elementary School and Professor of Education, State University of Iowa, Iowa City, Iowa.*

EARL HUDELSON, *Dean, College of Education and Professor of Education, West Virginia University, Morgantown, West Virginia*

FREDERICK J. KELLY, *Chief, Section of Colleges and Professional Schools, Office of Education, Department of the Interior, Washington, D. C.*

WILLIAM A. McCALL, *Professor of Education, Teachers College, Columbia University, New York City*

W. J. OSBURN, *2005 Waldeck Ave., Columbus, Ohio.*

SHELTON PHELPS, *President, Winthrop College, Rock Hill, S. C.*

DOUGLAS E. SCATES, *Director, Research and Statistics, Cincinnati Public Schools, Cincinnati, Ohio*

JESSE B. SEARS, *Professor of Education, Stanford University, California*

FRANCIS T. SPAULDING, *Associate Professor of Education, Harvard University, Cambridge, Massachusetts*

JOHN L. STENQUIST, *Director, Bureau of Educational Research, Department of Education and Lecturer, Johns Hopkins University, Baltimore, Maryland*

PERCIVAL M. SYMONDS, *Associate Professor of Education, Teachers College, Columbia University, New York City*

W. W. THEISEN, *Assistant Superintendent of Schools, Milwaukee, Wisconsin*

M. R. TRABUE, *Director, Occupational Research Program, United States Employment Service, Department of Labor, Washington, D. C.*

W. L. UHL, *Dean, School of Education, and Professor of Education, University of Washington, Seattle, Washington*

CARLETON W. WASHBURNE, *Superintendent of Schools, Winnetka Public Schools, Winnetka, Illinois*

GUY M. WHIPPLE, *Secretary, National Society for the Study of Education, Clifton, Massachusetts*

ELIZABETH L. WOODS, *Supervisor, Educational Research and Guidance, Los Angeles City Schools, Los Angeles, California*

G. A. YOAKAM, *Professor of Education and Director, Courses in Elementary Education, University of Pittsburgh, Pittsburgh, Pennsylvania*

THE JOURNAL OF EDUCATIONAL RESEARCH is published at Bloomington, Illinois, monthly, except June, July, and August by the Public School Publishing Company. The subscription price is $3.50; $3.50 in Canada or Mexico; and $3.80 in other foreign countries.

Subscriptions and all communications about reprints or other business matters should be addressed to the PUBLIC SCHOOL PUBLISHING COMPANY, at Bloomington, Illinois.

Address all communications on editorial matters to A. S. Barr, Department of Education, University of Wisconsin, Madison, Wisconsin.

JOURNAL *of* EDUCATIONAL RESEARCH

Volume XXVIII MARCH 1935 Number 7

A SYMPOSIUM ON THE EFFECTS OF MEASUREMENT ON INSTRUCTION

CLIFFORD WOODY AND OTHERS
University of Michigan

THE purpose of the symposium is to study the effects of measurement on instruction. A number of articles have appeared in recent months discussing one or another aspect of this problem. So important is the subject that it seemed worth while to bring together a statement of the views of a number of persons upon it. Of those invited to participate in this symposium ten have prepared papers stating their positions. These several statements follow.

THE EFFECTS OF MEASUREMENT ON INSTRUCTION

A. S. BARR
University of Wisconsin

THERE is, of course, much discussion today of the desirable and undesirable effects of measurement upon instruction. The effects of measurement upon instruction depends, it would seem, upon the kinds of measures used, by whom they are used, the uses to which they are put, and the skill with which they are used. While it is not uncommon for some educationalists to think of measurement as synonymous with objective tests, standardized or unstandardized, the term measurements is a much more inclusive term including such measuring instruments as tests, (recognition tests, recall tests, essay tests, etc.), rating scales, check lists, questionnaires and various sorts of mechanical instruments of measurements. Instruments of measurement may be standardized or unstandardized; they may be subjective or objective; they may be valid or invalid; and they may be reliable or unreliable. They may be used by teachers, principals, supervisors, superintendents, and other school officials. They may be used for purposes of research; for such administrative purposes as ability grouping, the rating of teachers and instructional surveys; for such supervisory purposes as the evaluation of in-

structional procedures, the search for causes of unsatisfactory pupil achievement, and the improvement of instruction; and for such teaching purposes as the discovery of individual differences, educational diagnosis and the measurement of pupil progress in the attainment of both the general and specific objectives of education. Any comprehensive survey of the effects of measurement on instruction would need, it seems, to take into consideration these and many other aspects of measurements.

In general, however, the good and evil of measurement can be attributed to either (a) the adequacy or inadequacy of the persons who make the measurements: their abilities, knowledges, skills, attitudes, ideals and interests in measurement, or (b) the adequacy or inadequacy of the instrument of measurement: their validity, reliability, objectivity, etc. While our instruments for measuring many of the products of instruction are wholly inadequate and while it will probably take years of painstaking effort to develop adequate instruments for the measurement of these more difficult to measure products of instruction the chief difficulty today in spite of all this, lies it seems, not so much in the inadequacy of our instruments of measurement as in the inadequacy of the persons who use them: their failure to choose measurements in terms of the specific purpose to be served; their use of instruments of measurement for purposes for which they were never intended; the neglect or non-measurement of products of instruction that cannot be measured objectively; their treatment of reliable instruments as if they were valid; etc. Of course, there are many aspects of measurement; and there is much to be learned, and while those who use measurements have learned much about how to use them, particularly during the last quarter of a century, there is yet very much more to be learned by all of us, before we can be said to have mastery of the art and the science of this very difficult educational function.

Aside from these general considerations the good and evil of measurement arises, it seems first of all, from the understanding or lack of understanding of the general purpose of measurement. The ultimate purpose of measurement is to supply the kinds of information that will terminate in the improvement of the products of learning. If the specific uses to which particular instruments of measurement are put do not promote this purpose, then their use constitutes a detriment and not an advantage in the program for improvement of instruction. Some criterion such as that given above should always be kept in mind in judging the effects of measurement on instruction.

Of the criticisms of measurement, if only a few are to be mentioned, it seems that when tests are chosen or imposed without due consideration for both the general and specific purposes of the instruction in progress that they may, because of their lack of curricular validity, give very erroneous impressions of the quality of instruction; the repeated application of measurement to certain of the products of instruction such as knowledges and skills to the neglect of others such as attitudes, ideals and habits of conduct, etc., may change the course of instruction along the channels emphasized in measurement, and the repeated use of the same sorts of instruments of measurement may result in desirable or undesirable standardizations. Of the many criticisms of measurement I refer to these three particularly because of the frequent reference to them in recent years. In all of these criticisms, however, the good and evil of measurement depends, it seems, as has already been said, not so much upon the inadequacies of the instruments of measurement as upon the inadequacies of those who measure.

Finally, I think that it ought to be said that educational progress is dependent upon the development of adequate instruments of measurement. As Jevons[1] said many years ago, the condition for advance in any science is dependent upon the invention of suitable instruments of measurement. "Forces hardly suspected to exist by one generation, are clearly recognized by the next and precisely measured by the third generation." So important is measurement that "the introduction of a new instrument often forms an epoch in the history of science." Those engaged in education have always measured; they are now measuring; and as far as I can see they will always measure. Measurement is so inherently tied up with the practice of education, with its evaluation and its improvement that I suspect as time goes on that we shall spend more time rather than less time in this very important educational function. Persons who appear to be opposed to measurement are really not opposed to measurement but opposed to some particular instrument or practice of measurement. In brief then, it would seem that the problem before the educationalist to-day is not that of eliminating measurement but that of developing better instruments of measurement and making those who use measurement more intelligent in its use.

[1]Jevons, W. S. *The Principles of Science* (London: The Macmillan Company, 1924), page 270.

The Effects of Measurement on Instruction

W. W. Theisen
Assistant Superintendent of Schools, Milwaukee

1. *Does measurement tend toward maintaining the status quo and toward stabilizing a low order of educational value?* That such an outcome has sometimes resulted, particularly where tests are used by untrained and unthinking individuals, must be admitted. The extent to which it results in given instances, however, probably depends upon the attitudes with which those in charge of testing approach the task and the previous training of the teachers whose children are tested. If measurement is limited to efforts to determine levels of academic attainment and teachers or pupils are rewarded in terms of the children's performances on tests, or if the results in any way are permitted to serve as a criterion of teaching success, such a result is quite probable. Unthinking individuals may easily become self-satisfied when results appear to confirm their own opinions of present practices. The current vogue of state-wide examinations to determine fitness for college entrance is quite likely to do just that, unless the responsible individuals in charge are careful to demand evidence of having reached other goals represented in the school program, as well as mere academic proficiency.

Unfortunately some of those who employ tests have been unschooled in their uses. Some misuse is therefore not surprising. Commercial publishers and even educators have been altogether too prone to encourage the use of tests when those who are to make use of them have not had proper training. Just as delicate instruments in the hands of a skillful surgeon are a very necessary part of his equipment but worthless tools in the hands of a novice, so too tests in the hands of the untrained may prove quite useless, if not baneful.

On the whole, however, far from contributing greatly to the *status quo*, measurement has been the direct and indirect cause of many of the educational advances of the past two decades. In fact, the roots of many, if not most, of the newer developments in education may be traced to the measurement movement. Measurement has stimulated thousands of teachers, principals, and educational leaders generally to think critically about results, to analyze objectives, and the means employed to reach them. It has led to a clearer understanding of aims and

purposes and marked improvement in methods and materials of instruction.

Far from tending to produce a feeling of satisfaction with the *status quo* in all teachers, measurement has tended rather to produce a state of *uneasiness*, which in turn has led to reflection and efforts to diagnose educational difficulties more fully. Measurement has been conducive toward attempts to discover the whys and wherefores of results and of ever present differences between groups and individuals. It has stimulated teachers and research workers to seek underlying causes. Unsatisfactory results have led to an examination of methods, curricula, and textbooks. This was followed by inquiries into the nature of intelligence and the possible influence of various extraneous factors on the child's school work. New curricula and materials of instruction have largely replaced the old. Scarcely a text for use in schools today is offered by publishers that does not acknowledge its indebtedness to the scientific movement in education. The forward march of measurement, as represented by gradual improvements in the validity and reliability of tests and in their diagnostic qualities, and the extension of their use to wider fields, has been such as to leave little opportunity for the teacher so inclined to long remain satisfied with the *status quo*. There is little more reason to believe that teachers will be satisfied with results as they are, than there is to believe that the owner of an outmoded automobile, even though it still functions with mechanical perfection, will long be satisfied with his old car. Only a teacher professionally dead could long be satisfied with the *status quo*. Such an individual would be satisfied even if no measures of achievement were applied. The cause of her complacency must be sought in sources other than measurement alone.

2. *Should measurement limit its efforts to informational aspects of subjects or should it embody all values stressed in the course of study, including such qualities as thinking, appreciation, and attitudes?* If test design and administration is limited to the field of academic achievement, the effect will be to minimize the importance of other types of educational results. To fail to include measures of other outcomes in the testing program, even though such measures may be somewhat crude and deficient in validity, is to lend support to the feeling that other outcomes are unimportant. We are reminded that anything which exists in some quantity may be measured, albeit crudely. School men are often not so much concerned about mere quantitative acquisitions of knowl-

edge as they are that children shall develop desirable attitudes, habits, ideals, interests, and methods of thinking. What evidence is there that the thinking habits important for modern social life, for example, are being developed? Most educators would probably hold that more important than to know whether a class has reached appropriate standards of proficiency in the mastery of the fundamental operations of arithmetic, is to discover the effectiveness of the thinking habits of the children, to identify the exact nature of their habitual errors in thinking, and to note the progress resulting from remedial instruction. They are interested to know whether the children are able to employ the processes connected with efficient thinking. To what extent are they able to relate causes and effects, to determine sequences, to note difference and similarities, to sift important from unimportant details, to organize data, and to arrive at proper deductions? These ends are generally accepted as necessary outcomes of instruction. Yet to restrict measurement to a determination of the amount of factual knowledge acquired is to sanction neglect to stress such outcomes. If the various objectives are to receive their appropriate teaching emphasis, curriculum makers should first determine upon a proper division of emphasis, after which the test designers and administrators should show by the tests employed that appropriate proficiency or progress is expected with respect to each objective. Tests of appreciations, attitudes, and interests so far developed have sought to substitute some degree of precision and definiteness for vague assumptions as to results, and have tended to bring about some semblance of organization in the efforts of teachers in this direction. There is great need for their extension.

We need not be greatly concerned, as some writers appear to be, about whether specific attitudes or interests which individual children may have developed are such as meet our approval or not. We are interested to know what social attitudes they are developing, whether they are being developed on the basis of adequate data, whether the children are developing worthwhile interests, and what ideals of a civic, moral, and social nature they are developing. Objective evidence that progress is being made is desired. When we know what is actually happening we may call upon the curriculum maker and the psychologist to prescribe instructional materials and procedures. Like the gadgets attached to the dash board of an airplane which enable the pilot to adjust the operation of the plane to the factors which enter into the success of the flight, so measurement must supply the educational gadgets

by which the teacher may chart her course. To omit the wind drift indicator or the altitude indicator would be to increase the hazards of flying and the uncertainty of the final result. Similarly to limit measurement efforts to certain aspects of the total teaching result only is to increase the uncertainty of the whole outcome.

3. *What is the ultimate effect of the measurement movement on the instructional process and what should be the ideal relationship between measurement and instruction?* To be most effective, testing must be accepted by the teacher as an important phase of the teaching cycle. This cycle is made up of three steps — planning, teaching, and observation of results. The purpose of testing is to increase the effectiveness of the third step, thereby making more intelligent planning of further teaching possible. Measurement properly conceived is an integral part of the complete teaching process and as such becomes an important agency in directing the choice of subject matter and methods of teaching. Almost any test, which has diagnostic value with reference to some phase of the learning or teaching processes when used intelligently by the teacher for the purpose of discovering the status of her children with reference to some specific accomplishment or lack of accomplishment, is quite likely to prove more valuable in its effect upon her teaching than the best tests when administered by others with only distributions and central tendencies reported to the teacher. However, in the latter case the testing may appear to be more nearly divorced from instruction than it really is. If tests are given for scientific purposes and the outcome is an improved instrument of measurement, or an improved course of study, or if new light has been shed on the relative effectiveness of different teaching procedures, the relationship between testing and teaching in general is evident. Both the immediate relationship between measurement and teaching when tests are employed by individual teachers and the long-range relationship as in more scientific uses need to be recognized. During the past two decades teaching in the schools of this country has undoubtedly benefited greatly from the intimate relationship that has existed between testing and teaching. The ultimate effect of the measurement movement so far observable has been to turn the educational searchlight with increasing intensity upon objectives, methods, and curriculum content, thereby forcing their re-examination, evaluation, and reorganization in terms of more clearly conceived educational and social purposes and more scientifically chosen procedures.

Tests: Aims, Procedure, and Types

Guy M. Wilson
School of Education, Boston University

Tests are too often referred to as if there could be but one aim, one procedure, and one type. This failure to analyze leads to generalizations, most of which are wrong except when applied very narrowly. An elementary analysis of aims, procedures, and types in testing will make this apparent.

Aims.—The three most prominent aims in testing are connected with teaching, clinical study, and administrative classification.

The teaching test is the test of chief concern. It should be given by the teacher, and she should be helped to adapt it for use as a teaching aid. Obviously, teaching tests must further the curricular aims of the subject, be in harmony with right methods of teaching the subject, and be so adjusted to the pupils tested and the work covered that success and encouragement follow. The teaching aim is the aim of greatest value; more about it directly.

Clinical testing aims at thorough study of individuals and permits much greater possibilities of success or failure. The child's possibilities must be explored, his needs revealed. The tests are not administered by the teacher, nor the scores reported to her. The pupil is not made painfully conscious of scores; he is merely encouraged to do his best. These tests, which do not influence teaching except through general recommendations, or specific recommendations for individual pupils, may cover any desired range. They are detrimental in the hands of the average teacher, although they may be excellent for clinical use, where the psychological reports just as a physician reports to a patient, in terms of general mental hygiene. In time, teachers may be professionally trained to the point of using clinical test materials, just as the physician is professionally trained (two years pre-medic, four years medical course, two years interne under directions; total eight years of training beyond high school).

Administrative tests such as the Stanford Achievement Tests, the Modern Achievement Tests, etc., aim at quick classification of pupils for administrative purposes. Such tests, properly used, may serve a useful purpose. They are usually a combination of clinical and administrative test features. They are not teaching tests and should not be permitted to influence teaching. As usually applied, such tests doubt-

less do more harm than good, since they are dominated by statistical aims rather than teaching-subject aims or pupil needs. The dominance of statistical aims leads to the inclusion of many test elements of too-great difficulty in order to smooth out the curves. From the standpoint of the subjects many of these same elements, and others are traditional and useless.

Most standardized tests have been developed from the administrative view-point, many of them by young Ph. D.'s who were long on statistics but short on teaching methods and subject aims. These various weaknesses combined to give us tests of exceedingly poor quality.

Gradually there is developing a change of administrative viewpoint which may lead naturally to disuse of standardized tests, or at any rate will lead to such handling of these tests as not to affect the teacher or teaching adversely. The old administrative view was that so much subject-matter had to be covered each year to avoid failure. The new view is that no child who works reasonably to capacity should fail. Under this view teaching and clinical tests serve much better. The old tyrant, "norm," is dethroned.

Procedures.—Procedures have been implied throughout the discussion of aims. They are properly determined by aims. If the administrative tests are discontinued or properly subordinated, the other tests will more easily adjust themselves to proper aims. Not easily, but *more* easily.

Let me illustrate. In developing my language error tests (World Book Company, 1923), I wanted to emphasize their value in improving teaching. I therefore used only a limited number of the two hundred four common language error elements available through previous studies[1]. I used twenty-eight elements in each of the three test forms, and provided for some overlap. This limited list of test items was accompanied by plans for motivating the child and improving the teaching. But being a youthful enthusiast in the movement, I worked out norms, reliability coefficients, etc., and included them in the manual.

"Reports to the author" on these tests during the next few years brought home to me more clearly the tyrannical nature of norms, and the ease with which the more valuable data on individual differences and needs, were neglected. Therefore, in preparing a second series of the Language Error Tests and a new manual[2], I tried harder to em-

[1]Wilson, G. M. "Locating the Language Errors of Children," *Elementary School Journal*, XXI (December, 1920), pp. 290-296.

[2]World Book Company, 1929.

phasize the teaching value of the tests, and was able to secure the co-operation of the editors to this end, although the "norms" had to be published in the manual in order to satisfy the clamour for them. Later studies on the number of language errors attaching to an individual pupil[3] have shown the correctness of limiting the elements in these tests to a small number, and providing a general plan for motivation, individual diagnosis, and teaching.

My conclusions are that teaching tests should serve good teaching, including motivation and individaul diagnosis; that clinical tests should be in the hands of competent clinicians and conform to good clinical practices; and that administrative tests of the usual type should be discontinued, or so managed as not to affect teaching and teacher morale adversely.

Types.—Types of tests referred to in the above discussion are (1) teaching, (2) clinical and (3) administrative. Classification on an entirely different basis is possible. From the standpoint of subject-matter, tests divide themselves into three other types: namely, drill, problem and appreciation, since these are the three main types of teaching. The project might be added to the above, but a project on the original Stimson-vocational basis, is, in its basic thinking, problem work. Therefore the procedures applicable to problem type testing apply equally to project, except that a vocational project such as raising an acre of potatoes may be checked not only as to basic problem thinking, but also from the standpoint of actual results.

Drill aims at making automatic certain responses, in the form of skills, habits or knowledge. If the aim of automatic response is to be secured, the amount must be small in order that success may be assured. The drill type of testing calls for inventory checking, and is satisfactory only on the basis of complete coverage. For example, a test in addition, assuming that addition is to be completely mastered, should cover the one hundred primary facts, the three hundred decade facts for column addition to $39 + 9$, and the eighty additional decade facts needed for carrying in multiplication to 9×9, total, 480 facts. In like manner, the process difficulties must be checked. These vary with children, but analysis of pupil difficulties has disclosed the main ones which obviously should be included in an inventory check.

[3]Unpublished. Particularly the studies of Alice Dunn, Boston University, School of Education.

Now, if we turn from the above specifications to arithmetic tests and submit them to complete analysis, we find that most of them fall short of the specifications[4]. They are, in fact, mostly general survey or administrative tests. It is doubtful if they serve a useful purpose. Their meticulous scalings to tenths of a grade, certainly rest on very insecure foundations.

On the other hand, the complete inventory test becomes also diagnostic and when placed in the hands of the teacher and properly used, can become not only a guide for individual needs, but also a source of motivation for the pupil.

The drill type tests are useful in the fundamentals of arithmetic, reading, language errors, spelling and, on an entirely different basis, in writing. They are undesirable in geography, history and the sciences, since they divert attention from the real aims. These later subjects call for associative thinking rather than memorization of facts. Problem type tests have not been successfully developed. The available tests in problem subjects check information on a drill basis and are therefore detrimental since they divert attention from legitimate aims. It is doubtful if satisfactory standardized tests for measuring pupil attainment can be developed in the problem subjects. The teaching approach in these subjects is constantly changing, not only on the basis of development of new data, but also from the standpoint of pupil interests and community background. Therefore the testing is done by the work itself and is not aided by the usual standardized test procedures[5].

Testing in the appreciation type of subjects, such as literature, music, art, and, for that matter, most of history, economics, physics, chemistry, biology, etc., has not been successfully done to date and doubtless will not be successfully done. This is due to the fact that appreciation is largely a personal matter, an affair of the emotions. Checking therefore is largely beside the mark and, as usually done, is sure to defeat proper aims. A scale for poetry appreciation such as Doctor Trabue developed[6] might be so used as to enable the teacher to understand the status of the child with reference to appreciation of poetry. Such scales could be valuable; but this is not the usual type of standardized test.

---

[4]Beattie, Louise. *Analysis of Standardized Tests in Arithmetic.* Masters Thesis, 1933, Boston University School of Education.

[5]The author's views on testing in problem subjects are more fully developed elsewhere. See Wilson and Hoke. *How to Measure*, chapters 23, 10, 12, 13.

[6]Abbott, Allen, and Trabue, M. R. "A Measure of Ability to Judge Poetry," *Teachers College Record*, March, 1921.

Criticisms of tests from the standpoints of aims, procedures and types, does not exhaust the possibilities of test-criticism. A test may be reasonably right in aim, procedure and type — for subject, and yet be very wrong because of unwise selection of test elements. Elements in an arithmetic test such as subtraction of compound numbers, subtracting a mixed number in eights from a decimal in thousands, solving an example in proportion, dividing a mixed number in 5ths by a mixed number in halves and asking that the answer be expressed as a decimal, extracting the square root of a five-place number, are very exceptional in adult usage and contain possibilities of great injury in tests for children.

Bad testing elements are not confined to arithmetic. In all subjects aims appear to be forgotten when standardized tests are being developed. In language we aim at sentence sense and correct usage; we test for "a noun in apposition," "an infinitive," "an adverbial clause," " a relative pronoun," "an intransitive verb," etc. In history and civics we claim to be making more intelligent voters, more loyal citizens, better Americans; we test by asking for fact data on Sir Francis Drake, Plymouth's military leader, Quebec, the Articles of Confederation, the Emancipation Proclamation, the XYZ affair, the Dred Scott Decision, Thomas Paine, etc. In geography we try to teach mutual interdependence and international good-will; but our tests call for facts on the Niagara River, the Indian Peninsula, Madrid, Australia, Brazil, London, etc. In short we forget that the technique of testing should be subordinated to teaching. The primary purpose of testing should be to aid teaching. Testing should be omitted unless it *is* good teaching. Standardized tests as they have developed in most subjects seem least likely to aid or further good teaching and should therefore be abandoned. With rare exception, they test on the lowest level.

The above statements are not too strong. All of the test items mentioned in the last two paragraphs are taken from tests used in a recent "nation-wide testing survey." Similar items are found in other recent achievement test batteries, and most standardized subject-tests.

The evident viewpoint of the writer is that the schools should serve the child and his proper development. Any testing which has a detrimental effect upon teaching therefore is not satisfactory. If clinical or administrative tests are given, they should be so separated from the teacher and teaching that they will not adversely affect them. The tests that are permitted to influence schoolroom practices should be teaching tests and should be administered by the teacher for the purposes of diagnosis, motivation and the direction of work.

Measurement and Instruction

L. W. Keeler
University of Michigan

In the past the results of instruction have been subjected to appraisals, usually carried forward on the basis of a scale of values which were partly objective but largely subjective in character. This scale was narrow in range because the concept of individual differences had not come into existence. With the rise of this concept and the development of modern science the growth of measurement in education was inevitable.

Instruction aims to bring the child into contact with environment in such fashion as will enable him to manipulate this environment with a minimum loss in time and energy. This process of self development takes place in terms of both quantitative and qualitative increment, and the quantitative aspect is open to measurement with the present instruments at our disposal. There is ground for hope that similar measurement of the qualitative aspects of development will become possible.

Subjective judgment is rarely capable of making very fine distinction in its appraisal of accomplishment. Careful training and much experience are needed to enable the instructor to appraise the various degrees of accomplishment of his class, upon the basis of subjective judgment, in terms of the individual differences existing among its members. The lack of such capacity usually results in undifferentiated instructional procedure; and too often produces a situation in which the needs of many individual pupils are neglected.

The creation of scientifically organized tests of general intelligence has made possible the determination of pupil capacity. The results obtained by the use of these instruments made imperative the organization of scientifically constructed instruments for the measurement of individual accomplishment. The result has been an approach to scientific description of pupil ability and accomplishment in objective terms. In consequence of this situation the effect upon instruction has been very pronounced.

Educational literature in the past two decades has shown the existence of two antithetical points of view regarding the value of measurement. One group decries this whole movement because instruments of measurement now available are inadequate to the job of measuring the entire effect of the instructional process. In opposition to this view

enthusiastic proponents of measurement hold that subjective judgment is influenced by too many halo effects, and that it is well to use even imperfect measures in an attempt to avoid errors inherent in such judgment.

In the early period of the development of scientific measurement the first group was undoubtedly encouraged in their view by unfounded claims made by enthusiastic supporters of the movement. They also were sure that the use of measurement magnified the importance of factual material at the expense of rational organization of the field in which this material falls. The question which naturally arises in this connection involves the extent to which rational organization can occur if the individual is poorly equipped with skills and knowledge essential to satisfactory procedure.

The second group is in danger of placing an undue emphasis on factual knowledge and skills which are not put to functional use. Instruction which is indoctrinated with this point of view is in danger of selecting for its attention so many individual trees that the student entirely loses sight of the forest. The dilemma thus raised can be resolved only by the development of instruments which shall measure both the qualitative and quantitative aspects of knowledge and skills. On the basis of what has been accomplished in the past we may confidently look to the future for such development.

The sound position to be taken between these extreme groups is that of acceptance of the inevitability of measurement and a recognition of the present fallability of our instruments and techniques. The generation who used the adze did not supinely await the rise of a generation which used the power saw and planer. They recognized the slivers left by the adze and took care to avoid them. We may be in the stage of educational alchemy but should hold to the justifiable expectation that our present gropings will result in the era of educational chemistry.

As the public schools are now organized we realize that no factor can enter into their procedures which does not affect instruction — the function of the school which is most vital to the development of its pupils. There is no doubt that the teacher is the vital agent of instruction and that there are many whose instructional procedures are well adapted to pupil development. There are many teachers, however, concerning whom this statement cannot be made and in their case im-

provement in instruction is essential to the development needs of pupils. What is the relation of measurement to this situation?

Psychological investigation has revealed many things concerned with instruction. It has shown the existence of various levels of pupil ability. It has also revealed variations in ability and performance at any one of the levels of general intelligence. It has brought to light the principles underlying the matter of difficulty inherent in various units of subject matter. It has also proven that pupil development is encouraged by a definite knowledge on the part of the learner of his rate of progress in achievement toward goals of his own acceptance. All of these factors must be recognized if instruction is to be adequate to its task.

If the measurement program includes the investigation of ability levels and levels of accomplishment, data are secured which will to a degree describe the efficiency of the instructional program and its administration. As a result of such program of measurement pupils, teachers, and administrators are not only given objective evidence of the status of instructional efficiency but they are also stimulated to search for possibilities of improvement. Is each student securing his fair opportunity for personal development? If not, investigation should be made of curriculum, instructional skills, personnel adjustment and coöperation, administrative procedures, and fitness of plant and adequacy of supplies. Careful and sensible measurement has resulted many times in stimulating personnel to secure a better understanding of the learning process as it applies to various types of knowledge and the laws relative to the development of skills. It has stimulated scientific curriculum construction and the selection of materials suited to the proper functioning of curriculum. It has often aided in the adoption of more efficient administrative procedures and has furnished administrative officers with such evidence of need as to enable them to secure from boards of education funds necessary for development and regeneration of school plant.

All in all it would seem that instruction and measurement should go hand in hand. Measurement divorced from instructional improvement is largely a waste of effort, but instruction without measurement can hardly hope to satisfactorily realize goals of achievement. In fact, goals set up without the service of adequate measurement are more nearly illustrations of wishful thinking than attainable reality.

Effect of Measurement on Instruction

Walter S. Monroe
University of Illinois

It is a common observation that in studying for an examination, students tend to make their objective the questions which they expect the instructor to ask. Their day-by-day study also is influenced by the types of questions asked during recitation periods. Furthermore, there is likely to be a close agreement between the objectives towards which the instruction is directed and the questions asked in class and on tests. In other words, a teacher's real objectives are reflected by the questions and other requests that he makes of his students.

When the tests administered are constructed by someone other than the teacher, that person tends to become the curriculum-maker, especially if the tests come from a source of influence. Since Dewey's experimental school focused our attention upon the curriculum, volumes have been written on the objectives towards which instruction should be directed, and there have been many conscientious attempts to apply the theories advocated in formulating objectives. During the same period, the measurement of results of instruction by objective tests has been popularized. Teachers, as well as administrators, have felt that a school was not respectable unless some use was made of standardized tests or of home-made objective tests. The result has been that the test-makers of the country have been the real curriculum-makers. This statement may appear to be an unjustifiable assertion, but the validity of the principle is apparent in the influence of college entrance examinations upon our secondary schools. If the movement for statewide and regional testing continues, we will soon find that the makers of the tests used have become the real curriculum-makers of our schools. Published courses of study and lists of objectives formulated by teachers will tend to become mere scraps of paper.

Granting that test-makers are the real curriculum-makers, some may ask, "Well, what of it?" The implied contention raises the question of the appropriateness of the objectives set by test-makers. Since the tests in question are objective, an important subordinate question relates to the type of instructional goals that are implied by such tests. One characteristic of objective tests is that the pupil does little or no writing. Hence, the goal of becoming able to organize ideas relative to a topic or question and to express them effectively is left out of the picture. A critical

examination of the various types of objective tests will reveal other differences between the goals implied by such tests and those to which we subscribe in theory.

It is a common observation that pupils who make high scores on tests frequently fail to apply what they appear to have learned. The explanation of this phenomenon is apparent when we note what the tests measure directly and what is required in application. The ability to spell a list of dictated words is not necessarily the same as the ability to spell the words employed in expressing ideas. The ability to identify statements that are true and those that are false is not the same as the ability to discuss a topic or to respond to some other type of thought question. A number of investigators have presented evidence to show a high correlation between the scores yielded by objective tests and measures of ability to respond to thought questions. This evidence, however, does not prove that the abilities measured by objective tests are the same as those that function in responding to thought questions. It merely shows that in certain situations the ability to respond to thought questions can be measured indirectly by measuring the ability to do something else. Under favorable conditions, such indirect measurement is only partially satisfactory and the more objective tests are employed to measure indirectly the ability to respond to thought questions, the less satisfactory the measures will be.

The limitations of space prohibit an elaboration of the argument, but the writer whose experience with objective tests extends over nearly a quarter of a century, welcomes the opportunity to register his conviction that the character of the tests employed in measuring the results of our teaching influences the instruction, and that the increasing use of objective tests is tending to nullify our attempts to change the curriculum.

The Effect of Measurement on Instruction

A. C. Krey
University of Minnesota

There may be differences of opinion regarding the meaning of "measurement." As commonly understood by teachers in whose classes various devices of "measurement" are used it is identified with the short answer form of testing students' knowledge, ability, or achievement. The teacher is vitally concerned in discovering the progress which the student is making in the particular field in which they are mutually concerned. If the psychologist is likewise interested and has a device which will measure that progress more accurately the teacher and the school will want to use it. The student is also intensely concerned in his rating and, since this rating depends upon his showing on whatever tests are used, he will adjust himself to such tests. Under these circumstances any "measurement" which is applied to students in the classroom inevitably affects instruction, both student and teacher assuming that they will be rated by this device. It is futile for the psychologist to insist that he has some other purpose in mind or that the "measurement" can be divorced from instruction. Only by keeping measurement out of the classroom and out of direct contact with the objects of instruction can any such separation be realized. The writer will assume first that measurement is concerned with the objects of instruction, that it will be used in the classroom and that it does affect instruction. The question which is left for discussion is whether the effect is good or bad, or necessarily good or bad.

Few persons not professionally concerned with measurement are in a position to distinguish between what is done by the elect in the field and what is done by all those who profess to be engaged in measurement. Is the hastily improvised test used by the classroom teacher to be regarded as measurement because it is cast in the short answer form? Is the so-called standardized test contrived by commercial test bureaus, textbook publishers, or even by research bureaus of school systems to be regarded as measurement? Or shall the term be confined to the devices employed by the psychologist in his laboratory alone? In the absence of any authoritative guide in the matter the average layman tends to regard them all as concerned in measurement and to reach his judgment as to measurement indiscriminately on the basis of all these instances of it with which he has come into contact.

It should not be surprising, therefore, to find our colleagues in subject-matter fields accepting the following quotations as a proper reflection upon measurement as they understand it.

". . . . He before the monarchs were the absolute rulers of the country, exercised his power in a spiritual way. This power was abused as the Pope became the authority on the justice of the people and tended to favor those with whom he wanted . . ."

"Professor, will the examination be objective or subjective?"

"I have been dismissed from my position (held for the past fifteen years) chiefly because my pupils have not scored satisfactory grades on the state board examinations in the past few years." (New-type form used five years).

"I have an objective examination for my course which I have used for seven years now." "I have developed an examination for my course in five different forms which I can use indefinitely."

They regard these quotations as symptomatic of definite defects. The first, an excerpt from a quiz paper of a senior student in college who is a candidate for a teacher's certificate in social studies, appears to them the logical outcome of continued use of the short-answer type of examination. They point to the correctness of the spelling as evidence that the student might have been trained to write much better English had not the excessive use of the short-answer examination deprived him of the necessary practice. The second quotation, which has become almost universal among our college students who have had much experience with the short-answer quiz, reflects their judgment that this type of examination calls for a very different kind of study. The third quotation reflects the same judgment and reveals an unnecessary and, in this case, an unjust dismissal of a teacher whose main fault was a lack of familiarity with recent developments in the use of short-answer tests. The fourth, in the minds of many of our colleagues, indicates a definite paralyzing effect upon educational and intellectual advance caused by a desire to achieve the utmost mechanical efficiency afforded by the short-answer form.

The writer, privileged by a much more intimate association than most of his colleagues in subject-matter fields with leaders in the field of measurement, realizes that the latter are as strongly opposed as the former to such unhappy results in the use of new type tests. The layman, using that term for all who are not psychologists active in measurement, requires more vigorous assurance on certain fundamental

questions. Is "measurement" to be identified exclusively with the short-answer form of test? May it not also include exercises in application of knowledge, laboratory observation and tests and even essay form questions? Must the process of standardization eliminate from measurement tests all items which are not in reach of every pupil? May it not also include a wide range of items some of which only the best pupils can be expected to know? Must a standardized test include only items which are definitely limited with a cold precision which in the social science field is true only of what is dead? May it not also include variables which probe dynamic knowledge? Is it essential to the scientific objectivity of measurement that its tests be constructed by persons who have little knowledge of the subject to be tested? Is it not possible to focus upon measurement of advanced subject-matter learning equally competent workers in both subject matter and measurement? Is statistical validity the sole criterion in determining the validity of a subject-matter test? Does not the significance of the items in the field of knowledge to be tested constitute at least an equally essential criterion of validity?

The layman, deriving his answers from the examples of measurement with which he has most frequent contact, even the so-called standardized tests now most widely used, can scarcely gain a favorable impression. And yet the writer feels certain that the majority of the psychologists with whom he has been closely associated during the past five or six years could and would answer every one of these questions in a way to satisfy most exacting laymen. In some instances they have already done so. Part of the difficulty in the way of a more complete rapprochement between subject matter and measurement scholars doubtless lies in their failure to achieve a common understanding of technical terminology. The writer's deficiencies in this respect have been most earnestly pointed out by Stalnaker.[1] Such difficulties, however, should not prove insurmountable, at least in matters of mutual concern.

If these questions can be satisfactorily answered and the difficulties overcome, measurement will continue definitely to affect instruction but its effect will be more beneficial. There will then be much less occasion for the fallacy, now so wide-spread, which assumes that a person whose test reveals some particular dexterity and interest is thereby assured a successful career as a surgeon, or that one whose test reveals

[1]Richardson, M. W., and Stalnaker, John M. "Comments on Achievement Examinations." *Journal Educational Research*, XXVIII (February, 1935), pp. 425-432.

a knowledge of miscellaneous and unrelated facts in history is thereby destined to become a successful historian or social scientist. A statistically correct effort to test a certain value does not mean that it has actually been tested in the opinion of subject-matter scholars, and the measurement of one element in an intricate complex of elements is not a test of the whole complex. If these qualifications become more widely recognized, there will be much less room for misunderstanding. And if, in the measurement of subject-matter learning, the test also seeks out the identical values which the teacher is trying to develop, the effect of measurement can only be to enhance, not hamper, instruction in that field. To these much desired ends both social scientists and psychologists may well devote continuous and sympathetic effort.

The Normal Relationship Between Measurement and Instruction

M. R. Trabue

Director, Occupational Research Program, U. S. Employment Service

The products sought in the educational industry are certain human qualities and characteristics which those in authority have selected because they believe them to be socially desirable. The various processes by which employees in this industry attempt to alter the amounts of these characteristics present in the raw materials coming to them are labeled "instruction." The procedures by which employees attempt to determine the amount of any characteristic possessed by an individual at any given time are labeled "measurement." The relationship of measurement to production is the same in the educational industry as in the automobile industry, the peach-growing industry, or any other planned effort at production. Measurements of the characteristics of the materials, in the "raw" state and at various stages before and after the "finished product" has been turned out and put into service, supply the basic data upon which intelligent planning of production methods must always be based.

No other industry of comparable size and importance attempts to get along with as few and as unreliable gauges and measuring devices as does education. A primary reason for this condition is that gauges and tests having the desired validity and reliability have only recently been developed. A great majority of the operatives in the educational industry still look upon the standard tests and gauges that have been

introduced either as direct reflections upon their own abilities to judge the characteristics of their pupils, or as harmless and temporary "fads" of the management.

It will probably require many years of patient and painstaking work before the average school executive and his planning division will be able to employ efficiently these precise records of standard objective measures as the basic data for the development of educational policies and instructional programs. There can be no doubt, however, that this will be the ultimate and normal relationship between measurements and production in education, as it now is in every other well developed industry.

Many of the questions raised concerning measurements in education may be answered by very simple analogies. When one climbs into a modern motor car or airplane with the intention of going to a particular place, he finds before him a complex array of gauges and indicators. The records that have been provided by these measuring devices on previous trips may assist the driver to determine his course intelligently. He may recall, for example, that his speedometer registers 46 miles by Highway Number 12, but only 36 miles by Highway Number 83. The use of this speedometer does not of itself change the course the driver will take, but the records it has supplied may be used to help him choose his course intelligently. If the shorter road is so mountainous that it is known from previous records to consume more oil and gas, the difference of four miles in distance may be entirely disregarded. If the driver is wealthy and interested chiefly in speed, he may consider only the one measurement of distance. If he is frugal and wishes to save his car and fuel, he will probably consider other measurements as well as that of distance. If he is under the direction of a vigorous supervisor, he will undoubtedly make his decision in terms of what he believes to be the supervisor's desires, even though he knows that the results will be uneconomical and inefficient. Measurements do not of themselves control our outcomes, but they do provide the data by which we may choose and achieve outcomes intelligently and efficiently.

Perhaps a more complex enterprise would be recognized as more fully analogous to education. Industrial recovery is now being attempted in this country by many different types of effort: social, political, educational, economic, and so on. The present status of industry is being measured by many different types of indices. The inter-

pretation of the significance of any particular index varies with the politics, economic status, personal ambitions, and social vision of the person reading the index.

Intelligent social planning, whether it looks toward increasing the well-being of the many and reducing the profits of the fortunate few, or whether it aims toward the extension of economic advantages to a few and still greater reduction in the well-being of the many, — whatever purposes and objectives may be in mind — intelligent planning cannot be done without the valid and reliable data supplied by the most objective measuring devices available. The measuring instruments themselves do not strengthen either the *status quo* or the "evolving social order," however vigorously the advocates of either program may use these measurements in propaganda for their cause.

Responsibility for changes in instruction rests, and always must rest, upon those who interpret data rather than upon the measuring devices by which the data were obtained. It is true that many measuring devices are deficient in validity and in reliability, but it is the responsibility of those who make plans and programs to know the relative significance of their various sources of data. The mariner's compass does not often point directly north, but it would be absurd to try to choose a course upon the ocean without the use of such an instrument. No one would trust the captain of a ship who insisted on throwing overboard his compass and other measuring devices just because they were not perfect. The compass itself does not change the course of the ship, but it does provide basic data by which the ship may be more intelligently steered. If some helmsman guides the ship straight north "because the compass points that way," he should be taught how to use the compass rather than required to steer without one. Similarly, the teacher who assumes that all her pupils should score "high" on all the tests given by her supervisor should be taught how to use tests intelligently — if, indeed, the supervisor himself knows their proper function. We need in education many more objective tests and much more intelligent use of them. We should rid the profession of inefficient teachers and incompetent supervisors, perhaps, but we must make more intelligent use of a larger number and a greater variety of objective measuring instruments in determining the effectiveness of our efforts and in planning our programs of instruction.

The Effect of Measurement on Instruction

John M. Stalnaker and M. W. Richardson
University of Chicago

It is an interesting commentary on the field of education that a symposium on this topic should be suggested. What results would be brought forth by a similar symposium on the effects of measurements on medicine, on physics, on engineering? The effect of measurement on any field is to make possible quantitative thinking in that field—in short, to make possible the development of a science. No one at this stage of the development of education is reactionary enough to suggest that measurement in education is a bad thing. One may criticize the emphasis which certain types of tests have been thought to make; one may suggest that value would result from concentrating more on measurement of other types, but one does not suggest that measurement should be done away with as far as the problems of instruction are concerned. Fundamental changes in instruction will continue to be based on the results of measurements. In fact, the field of education will gain recognition and respect in proportion to its sound use of measurement.

In the following discussion, we shall limit ourselves to the influences of *achievement examinations* on instruction. The discussion should be considered in the light of the purpose of achievement examinations—the measurement of absolute attainment of certain educational objectives for the purpose of academic certification. We shall make three general points about which considerable discussion has been aroused.

1. Achievement examinations have a desirable influence on instruction in that they demand that an objective of instruction be stated in terms of student response. Any objective so stated can be measured. Examinations are thus causing a definite clarification of objectives and their restatement in meaningful terms.

An examination is an academic tool, not an end in itself. Not even the most enthusiastic proponent of the widespread survey examinations expects that the particular forms developed are to be regarded as permanent irrespective of reform in curricula. In the ultimate analysis, examinations derive their content and form from educational objectives, which in turn reflect that shifting field of opinion known as social philosophy. One cannot prove or disprove by any technical means at our disposal that an educational objective is or is not valid. One consideration only is essential for our purposes: having decided on some

objective or having had some objective given to him, the teacher is forced to define this objective by expressing it in terms of what the student must do about it. The student may be required to memorize, to apply principles to new situations, to select, organize, evaluate, etc. The unescapable point is that the things the student does must serve to define the actual objective in any teaching situation; if those specified elements of student behavior represent the objective imperfectly, so must a measuring device which is based on them. The failure of examinations to represent higher objectives is perhaps due to haziness of objectives when translated into behavior expected of the student. One suspects that unrealizable hopes motivate the criticism that examinations reflect (and tends to stamp into educational procedure) mere acquisition of facts and other "low-order" elements. The crux of the matter is that most objectives have been resolved into about the same set of student operations and what is done (or not done) in teaching and curriculum-making is reflected in the measuring instrument. If the objective can be stated in teachable elements, student achievement can be measured by sampling those same elements. The simple requirement of defining the objectives in terms of student behavior is not arbitrarily imposed, but in the nature of the case inevitable. It furnishes a sane corrective for certain nebulosities in modern education. It does not prevent a teacher from trying to attain any objective he wishes to formulate. If he fails, it is not legitimate to blame it on the examination.

2. The contention that tests do not measure "higher" values is a negligible one. If these higher values can be agreed upon and can be stated in terms of definite procedures, they can be measured.

Some educators imply that the present emphasis on tests tends to minimize the "higher" aspects of education — the development of appreciations, attitudes, interests, ideals, habits of initiative and independent thinking, etc. But these are merely hoped-for outcomes of instruction; we really have no techniques for getting them across. It is a common pedagogic fault to let good intentions take the place of rigorous analysis of teaching operations: the comfortable feeling the teacher gets from hearing himself talk causes him to be less critical of teaching operations than of tests. Some courses of instruction will not withstand the scrutiny of careful measurement of results; nothing could be more natural than for the teachers of such courses to aver that examinations cannot measure the higher elements inculcated by these courses. These courses, like football and the Boy Scouts develop charac-

ter, ideals, personality, and the like, it is claimed; if one tries to pin these teachers down to their methods of developing character and appreciation, one will get unsatisfactory answers. Attendance in their classrooms, however, will show that the examination reflects just the sort of thing they are doing.

It has been urged that certain attitudes be consciously and systematically taught.[1] We should dislike to see any body of educators or politicians set up definite attitudes to be systematically taught to students as is done in Germany, Italy, and Russia. A group of social scientists at the University of Chicago has expressed itself on this issue as follows:

It is our opinion that the group of subjects called the social sciences should be concerned with the limited responsibility of attempting to understand and to explain the facts and relationships arising in an extremely complex natural phenomenon, human society and human social organization. We do not believe that the social sciences on any level should be used consciously for purposes of indoctrination or for the inculcation of standards, values, attitudes, no matter how generally accepted or apparently desirable. It is our opinion that the greatest educational advantages will arise from the attempt to keep or to make the social sciences scientific, and to leave for other subjects and other social agencies, the family, the church, the political party, etc., all activities dealing with the beautiful and the good as purpose and value to be inculcated for reasons of social integration and social management.

One does not entirely avoid the issue by asserting that only such general non-controversial attitudes as "the scientific attitude" are to be taught. Perhaps there is no such entity as "scientific attitude"; one might be rigidly scientific about cosmic rays and ridiculous about religion, politics, or economics. Tests can be made to measure, somewhat crudely to be sure, attitudes, interests, even the emotional response to a sonnet. The evaluative type of appreciation we feel should be measured, but we wonder whether the positive reaction of a student's autonomic nervous system to "A flower in the crannied wall" should count for or against him in academic accrediting. We doubt that progressive educators would seriously advocate the use of measured emotions as basic accrediting data, even though the necessary techniques for measurement are available.

3. Achievement examinations have a desirable influence on instruction in that they hold up absolute standards and thus call attention to the individual needs.

[1] There is evidence of changes of student attitude (in a liberal direction) with academic training. This incidental effect has operated presumably without any definite objectives.

It has become the fashion to deplore the effects on the individual of group or mass instruction; and it has been suggested that examinations constitute a vicious part of the regimentation of education. It is argued that the creative personality can be developed only by individualized instruction, with no deadly educational lockstep and no measurement by absolute standards. It has been suggested that if any measurement is to be done, the progress the individual has made should be measured. According to this system, if a student progresses uniformly, he gets the same mark at all times; moreover, he will get the same mark if he begins at the lowest conceivable level and makes progress to a final point much below the average of his age or of the citizenry in general. One would not be surprised to find some ingenious and sentimental person suggesting the use of the rate of change in the rate of progress. Then a student's marks would go up when his rate of attainment itself increased, and his personality would not suffer the blighting influence of a steady attainment. However, it is better to recognize that one necessary requirement for educational measurement is the use of the same scale for all individuals under consideration. The use of a standard foot-rule to measure the height of all persons is not reducing the average height of the race; not even if some of the replications of the standard foot-rule are faulty. Widespread psychological and educational measurements have called attention to the necessity of providing for individual needs; it would be foolish to minimize the use of devices which throw this problem into such strong relief. We do not provide for such differences by such monkey-business as accrediting students according to their own progress in time. Teaching or diagnostic tests might operate in this fashion, but achievement examinations have quite a different purpose.

* * * *

We have attempted to show that the achievement examinations reflects the course content as taught. If educational objectives are stated in terms of student behavior, additional types of tests can be devised to fit the requirements. If the emotional concomitants of learning are to be measured, they too can be. Such terms as attitudes, ideals, habits of clear thinking, ability to organize, etc. we have suggested are in general too nebulous to find their way into actual classroom procedures. If such objectives are defined and stated in meaningful terms, they can be measured. The examination clarifies objectives and forces out ob-

jectives that are practically impossible. Through the coöperation of test technicians and persons thoroughly grounded in the objectives of a course — as expressed operatively — adequate measures will be made, and once made they can be published and widely used. Tests are not static things. They are changing and improving and they will doubtless continue to reflect the actual classroom procedures.

The Effect of Measurement Upon Instruction

Harl R. Douglass
University of Minnesota

When electricity, hogs, wind-velocity, areas, distance, or liquids are measured, the act of measurement ordinarily has no effect upon the object or force measured. No matter what its weight, its loss, or its gain, the hog will weigh no less, no more, nor gain any desire to weigh less or more, by virtue of its having been weighed. The measurement of achievement of human beings is quite another thing. Because this is so, one cannot consider the techniques of measurement apart from their probable relations to learning and factors directing and conditioning learning.

In the following paragraphs are enumerated a few of these relationships which need be kept in mind in evaluating procedures and programs of measurement. Some of these have been ignored or given so little attention by some specialists in measurement and many novitiate evangelists that the development of sane and effective measurement has been seriously handicapped, and through their failure to survey these relationships, there is being enacted a chapter in the history of measurement that will be no credit to the movement.

1. Much more care should be used in defining and keeping in mind what is measured by a given test. A test for American literature, for example, may measure certain kinds of information, or certain skills in analysis and criticism, or certain tastes and appreciations, or certain interests, or certain complex abilities such as writing sound, interesting, and structurally impeccable criticisms or some combination of these outcomes.

Much loose thinking unworthy of scientists and educators may be traced to the fact that they have employed scores on achievement tests to represent achievements, abilities, or possession of other outcomes of instruction and study which the test scores do not represent, or represent badly. Experiments in educational organization or method

have been interpreted as if the test scores involved measured progress or growth toward all of the objectives of instruction and in such a manner as to constitute a representative balance of the various objectives.

2. It must be recognized that in few if any fields are there tests which measure progress or growth toward all the important objectives of instruction in that field. The interpretation of test scores, therefore, in the evaluation of teaching procedures, in the status of pupils, classes, or schools, in educational experimentation and in all situations must be developed in the light of this restriction.

3. The unmeasured outcomes of study and teaching may or may not be in high degree of correlation with the measured outcomes. It is not wise to assume that if the test measures the acquisition of information well, it probably also measures the ability to appreciate and the possession of any given ideals or attitudes as well. This begging the question is not likely to lead to truth and is dangerous and an unworthy avoidance of real difficulty.

4. The development of means of measurement must be stimulated to continue to expand until useful means of measuring growth in all the more important outcomes of study and instruction have been developed. To pretend that the measurement of ideals, attitudes, interests, or any of the affective outcomes of learning or the more complex skills lie outside the field of measurement of educational products is a premature surrender, implying that these types of outcomes are after all illegitimate and that the school should confine its efforts to storing children with factual information and drilling them in the simpler subject matter skills. Such a solution of the difficulty is not a solution at all but a pretense that the problem does not exist.

5. The importance of test taking as a learning activity is not significant. The criticisms that the use of objective tests is bad training since they encourage guessing are not convincing. The taking of tests need not be thought of as a learning procedure at all. In any ordinary situation the effect of a few hours spent in taking a test upon the pupil's mastery of the course, his habits of thinking, or any other educational outcome is certain to be relatively negligible. The testing itself is for the purpose of measurement, not for the purpose of teaching or learning. The foregoing statement should not be interpreted to mean that the test is not useful for instructional purposes. The test may be employed

with great effectiveness in motivating learning activities and in diagnosis for the guidance of learning and teaching.

6. The effect of the type of test upon the study attitudes and procedures cannot be ignored. Studies[1] and testimony of students indicate quite clearly that the methods of study and the relative emphasis placed by students upon various types of instructional materials are definitely influenced by the type of test employed.

7. Certain types of use and interpretation of test scores wield a tremendous influence upon teachers, pupils, principals, supervisors, and superintendents. Wherever the same tests have been given annually to a number of schools, or classes within a school system, the universal and inevitable tendency has been for instruction to be directed towards doing well on the test. This motivating influence may be a powerful asset or a tragic liability. To direct instruction most effectively toward doing well on a test which fails to emphasize one or more of the more important objectives of instruction in a given subject is just another way of saying that the aims and materials of instruction have been unfortunately distorted.

These conditions are peculiarly aggravated when the scores of pupils of different teachers or schools are compared and the reputation of instructors or institutions thereof related to the standings on such tests. No influence could be more inimical to progress in the educational philosophy of teachers and to the improvement of courses of study.

8. The use of tests by supervisors for the purpose of evaluating teaching efficiency is not only subject to many pitfalls in arriving at fair measures of relative efficiency, but the effects upon instruction can be wholesome only when the tests used measure all progress toward all the objectives of the course and in something like the proportion of their relative importance.

9. If and when tests may be employed which do measure in due proportion or separately all the important outcomes, the stimulating influence upon teachers, pupil, principal, and superintendent is certain to be not only considerable but extremely desirable.

10. Placing emphasis by means of contests or otherwise upon scores made on tests prepared by some agency other than that responsible for courses of study construction or supervision has two serious lim-

[1] Douglass, Harl R., and Tallmadge, Margaret. "How College Students Study for New Type Examinations," *School and Society*, XXXIX (March 10, 1934), pp. 317-318.

itations: (a) it is bad supervisory organization to have teachers in actual practice attempting to follow two leaderships in their selection, emphasis upon, and teaching of instructional materials, and (b) it is questionable policy to surrender leadership by local authorities, to an outside agency over which the local units have little or no control. Once having participated for a period of years in such a program it becomes very difficult to withdraw.

In final analysis, the character and force of the influence of testing upon instruction is not only a matter of the nature of the test but even more a matter of the use and importance allotted to the test scores.

Coöperative Achievement Testing

E. F. Lindquist
State University of Iowa

The past few years have witnessed a rapid increase in the number and scope of wide-scale educational achievement testing programs. These programs have appeared in a wide variety of forms. In some instances, they have been sponsored by state-supported educational institutions or by other responsible professional educational organizations, and have been conducted on a coöperative, non-profit or subsidized basis as bona-fide service agencies for the schools in a limited geographical area. At the other extreme, they appear in some cases to have been no more than mere test-selling devices, or competitions conducted simply for the sake of competition. In some of these programs, the examinations provided have been of excellent quality; in others, they have been noteworthy only as examples of poor test construction. The fears justifiably aroused by some of these instances, strengthened by previously established prejudices against objective testing in general, have unfortunately not been restricted to their sources. As a result, the desirability of its influence upon instruction has been made a general issue in relation to the whole coöperative testing movement. It is the interest recently shown in this issue, rather than in the broader question of the influence upon instruction of measurement in general, that appears to have been the immediate reason for this symposium. It is with this issue, therefore, that this article will be exclusively concerned.

It may be noted first of all that the generalized form in which the issue has thus far been presented is in itself not conducive to con-

structive consideration. Existing testing programs are too diversified — in avowed purposes and in underlying motives, in administrative organization, in nature and quality of the examinations used, and in observable consequences — to permit any safe generalizations. Coöperative testing both involves real dangers of undesirable consequences and presents significant possibilities for the improvement of learning conditions. Numerous instances can be found in which these consequences have not been successfully avoided and in which these possibilities have not been fully realized. There are ways, however, of minimizing these dangers and of realizing these possibilities more fully, and in some programs significant and reasonably successful attempts are being made to achieve these ends. In this situation, each specific program, or type of program, must be evaluated on its own merits. Any attempts at generalization are likely only to conceal the important fact of wide qualitative differences between existing programs and to handicap and interfere with the attempts that are being made to improve them.

In this article, therefore, the writer will be concerned with only one type of testing program, as exemplified by the Iowa Every-Pupil Testing Program, and in terms of this illustration will discuss certain of the possibilities presented and of the dangers involved in coöperative testing programs of this type.

The Iowa Every-Pupil Testing Program is similar in many respects to most of the outstanding coöperative testing programs now being conducted elsewhere in the country. It is sponsored by the College of Education and the Extension Division of the University of Iowa, and is conducted on a non-profit basis for the high schools of the state. It provides each year a new set of comprehensive objective achievement examinations which are administered at the end of the year, at the same time and under the same conditions, to *every* pupil regularly enrolled in the subjects tested in each of the participating schools. The papers are scored in the local school systems and the results reported immediately to the director of the program. A summary report of the entire program containing the norms on the tests, is then prepared and distributed to the schools about two weeks after the date on which the tests were administered. Confidential reports are also sent to each school independently, showing how its average performance on each test compared with those in all other schools participating. Participation in the program is completely voluntary. Over 250 high schools,

enrolling approximately 50,000 pupils, have been involved in each of the recent annual programs.

THE ADVANTAGES OF COÖPERATIVE TESTING

Space will permit only a brief and partial review of the advantages to be gained through this type of coöperative organization in achievement testing. These advantages involve:

1. *The provision of better tests.*—The construction of a good objective achievement examination is a highly technical and time-consuming process. Very few high schools, even in the larger communities, have the facilities, the time, or talent that are required. In a wide-scale testing program, however, where the same examination is to be used by thousands rather than by tens of pupils, it becomes possible and practicable to secure the services of experienced and highly trained specialists in test construction in each field, to secure criticisms and suggestions from a number of competent teachers and subject matter authorities, and to submit preliminary materials for experimental try-out. The fact that a new set of tests is provided each year precludes the possibility of coaching on specific items, permits continuous adaptation to changing curriculum practices, and makes possible the immediate application of the most recent developments in testing techniques. The fact that the tests are constructed under central supervision makes possible a consistency in quality and form of the tests and hence a degree of comparability from subject to subject in the measures obtained that rarely characterizes independently constructed tests. Finally, the restricted scope of the program permits specific adaptation to the curriculum practices of the schools in the limited geographical area involved.

2. *The provision of better norms.*—Under the so-called "independent" procedure in achievement testing, the local school system in, for example, Midtown, Iowa, would obtain for itself the specific tests that it might require, usually from a number of different sources, and would administer these tests at the time and under the conditions which best suited its own convenience. If standardized tests were used, the norms for these tests would have been established in advance — never less than one year, and frequently several years before the date of their administration in Midtown. These norms, therefore, could not take into consideration the results obtained in Midtown or in any other schools administering the tests this year. Ordinarily,

the specific schools in which the norms had been established would not be known to the Midtown superintendent. In many cases, they would have been located in distant states, and may or may not have been operating under conditions similar to those in Iowa. Furthermore, since the norms for different tests would have been established at different times and for different groups of schools, they could not safely be considered as highly comparable from one test to another.

Under the so-called "coöperative" procedure, on the other hand, the Midtown school would coöperate with a large number of other Iowa schools, all of which had agreed, to their mutual advantage, to administer the *same* tests on the *same* day and under the *same* conditions to *all* of their pupils. The norms for these tests would then be established *subsequent* to their administration rather than in advance of it, and would be based upon the results from *all* the schools which had administered the tests *this* year rather than upon a much smaller number of schools which had administered the tests last year or several years ago. All test results would thus be utilized with maximum effectiveness in establishing the norms provided, and the norms would be strictly up-to-date. Furthermore, they would be highly comparable from test to test, since all norms would have been established at the same time and under the same conditions for exactly the *same* group of pupils and schools. The norms would also be relatively high in reliability and representativeness because of the large numbers of schools involved and because *every* pupil had been tested in each participating school. Finally, the norms would be more concretely *meaningful* to the Midtown superintendent, since they would have been established for schools that are known to him and that are all operating under similar conditions — i. e., they would be *Iowa* norms.

3. *Low cost.*—Due to the non-profit nature of the project and to the economies made possible by large-scale production, it is apparent that all materials and services may be made available to the participating schools at a cost much lower than that at which they might otherwise be secured.

4. *The utilization of test results in research.* A centralized testing program, collecting as it does every year a tremendous amount of test data and related information from hundreds of school systems, from thousands of teachers and from tens of thousands of pupils, presents unique opportunities for educational research, and makes possible the investigation of many hitherto untouched educational problems at

very little additional cost. The utilization of this valuable by-product of coöperative testing in no way interferes with its fundamental purposes and provides an unequalled opportunity for studying the factors affecting learning, and for investigating problems of a type which are beyond the facilities of the ordinary research agencies.

Space is not available here for any discussion of the many specific uses which may be made of the examinations provided by programs of this type or of the new uses and the increased effectiveness in relation to all purposes which are made possible by the characteristics just noted. It is hoped that certain of these uses, particularly those in educational guidance, will be adequately considered by the other contributors to this symposium.

DANGERS INVOLVED IN COÖPERATIVE TESTING

The dangers involved in this type of testing may, at the risk of exaggeration, be identified by reviewing some of the criticisms which have, from time to time, been directed against it. These criticisms have included the contentions that, in their desire to secure higher test results, the teachers in the participating schools will "point" their instructions toward the examinations used; that they will become "drill masters," primarily concerned with the rote learning of facts culled from previous examinations; that they will seriously neglect the less tangible but more important outcomes of instruction which are not measured by the examinations; and that they will secure spuriously high standing on the tests through intensive last-minute cramming procedures. It has been claimed consequently that the tests used will "dominate" the high school curriculum and that the coöperative testing program is but another device for taking from the high schools the right to determine their own curriculum policies.

That wide-scale coöperative testing involves real dangers of the type indicated is not to be denied. On the contrary, proponents of coöperative testing should be particularly active in drawing the attention of test users to these dangers, since there is perhaps no more effective way of minimizing the dangers themselves than by making test users and test builders in general more fully aware of their existence. It is the opinion of the writer, however, that these dangers have been unduly exaggerated and that the undesirable consequences of which the critics have been so fearful have not, in fact, actually characterized many of the existing programs to any serious degree.

Certain of the factors which have contributed to these exaggerations are readily apparent. One of these is the still prevailing prejudice against the objective or "new-type" test — a prejudice which has remained blind to the many significant improvements recently made in objective testing techniques. Another is the strong prejudice against external testing agencies in general which resulted from the earlier unfortunate experiences of the high schools with such examining bodies as the college entrance examination boards, the New York State Board of Regents, and other state examination boards.

The afore-mentioned fear of "domination," for instance, appears to be largely a survival of these latter prejudices. It should be noted, however, that programs like the Iowa Every-Pupil Testing Program are being conducted not, as in these earlier instances, for the colleges by persons primarily interested in retaining the traditional college curriculum, but for and in the interests of the high schools by persons now engaged and primarily interested in secondary education. Furthermore, this type of program does not, as did those earlier agencies, dictate in any way the specific subjects in which each pupil must be tested. Participation in all of these programs is completely voluntary. There is nothing in them that can conceivably coerce the schools to offer subjects which they themselves do not wish to offer, or that will induce them to continue teaching subjects that they would like to eliminate. In programs of this type, each school is perfectly free to omit the use of any test that does not seem to conform to its own objectives of instruction. There are then very significant differences between this type of program and those which originally gave rise to the fear of "domination." Since programs of this type are being conducted in the interests of the high schools, the use of the term "domination" (with its undesirable connotations) in the attempt to discredit these programs is distinctly unfair to the motives for which they are conducted.

It cannot be too strongly emphasized that it is not the proper function of any testing program to attempt to influence curriculum content in any direction, and it is certainly not the intention of the sponsors of existing programs to use them in that fashion. The function of a general achievement test is to discover differences between pupils in the extent to which they have achieved that which they are actually being given to achieve, or, in other words, to discover how well the pupils are able to understand and make use of that which they

are now actually being taught. It is *not* the function of such tests to discover how that which the schools are now teaching compares with that which anyone else believes they ought to be teaching, or to rank the schools using these tests in the order in which the local curricula conform to any arbitrary ideal curriculum. In building the tests used in the Iowa Every-Pupil Testing Program, for example, the test authors have analyzed each year as carefully as possible the content of the commonly used textbooks and courses of study and have tried to limit the test items to the elements and concepts which are *now* universally presented in Iowa high schools. In a very real sense it is the test authors who are "dominated" or restricted by present curriculum practices rather than the reverse.

In asserting that it is neither the proper function nor the intended purpose of coöperative testing programs to influence the content and methods of instruction, the writer does not mean to deny that, regardless of any intentions, such influences nevertheless may be and often are exerted, or that these unintentional influences are in some instances definitely undesirable. Ideally, the tests provided by these programs should be considered by the teachers purely as *measuring* instruments and as nothing else. They should no more be considered as standards of curriculum content than weighing scales or yardsticks are to be considered as standards of physical perfection. Actually, however, teachers in general are not yet ready to view the tests only in this light, and cannot be expected to do so as long as their administrative superiors continue to use test results alone as measures of their teaching effectiveness. In the construction of tests to be used in these programs, then, great care should be taken to eliminate any characteristics (even though wholly unobjectionable or even desirable as far as the *measurement* function alone is concerned) which might unintentionally influence teaching in a wrong direction.

Special consideration, in this connection, should be given to the danger of encouraging mechanization of teaching. There is considerable evidence that high school instruction, even where entirely uninfluenced by testing programs, is in general overly dependent upon the textbooks; that mechanical rote procedures tend to predominate in learning and in recitation; and that what pupils actually acquire as a result of instruction frequently consists, in large measure, of a mass of more or less meaningless verbalizations. Pupils tend to depend upon the recall of "pat" expressions taken from the textbook and upon words

and phrases which have been memorized with no appreciation of their meaning or significance, rather than upon a reasoned understanding of the concepts involved. These undesirable practices in teaching long antedated the objective testing and coöperative testing movements and clearly cannot be attributed to them. It must be recognized, however, that depending upon the nature of the examinations employed and the uses made of them, coöperative testing programs may unintentionally perpetuate or even accentuate these undesirable practices.

To avoid these consequences, the tests used in coöperative testing programs should be constructed with the conscious and deliberate purpose to *defeat* rote learning of, or mechanical and verbal drill upon, poorly selected and isolated information. In them, an attempt must be made to avoid the use of stereotyped textbook language; to reduce to a minimum the proportion of "pat" questions of the familiar "name, identify, define, and describe" or "who, what, when, and where" types, the answers to which are likely to have been memorized without understanding by the pupil; to call for *applications* rather than *statements* of laws and principles, and in so doing to use novel illustrations, new settings, and unfamiliar approaches; to require the *interpretation* of real problem situations, rather than the mere repetition or recognition of words; to place the emphasis upon *understanding* of interpretative ideas and ideas of relationship, rather than upon verbal learning of isolated items of descriptive information; and in general to stress the *functional* values in what has been learned, rather than the "subject matter" itself.

The tests used in existing programs may in some cases be conspicuous only for the absence of these characteristics, and in such cases it is quite likely that real damage may have been done. It is quite possible, however, to construct objective tests that will exhibit these desirable characteristics to a very satisfactory degree. If such tests are used in a coöperative testing program, and if teachers and pupils are made aware of these characteristics of the tests, then it is hardly conceivable that they will accentuate the practices already noted. On the contrary, they should then exercise a very desirable and wholesome influence — but an influence, it should be noted, which is concerned with the *quality* rather than with the *content* of instruction. Without attempting to effect changes in or to "dictate" to the schools the basic content of current instruction, such tests can do much to motivate more effective teaching of that content, and can place a very desirable premium in achievement upon understanding of and ability to use the "subject matter" acquired.

The provision of tests of even the highest quality, however, is not itself a sufficient assurance that the influence exerted by a coöperative testing program will be wholly desirable. Before these programs can realize their possibilities, a great deal must be done to educate teachers and school administrators to a more intelligent use of the materials provided. In the first place, test users in general must be led to a more adequate understanding of the true nature and functions of general achievement tests, and to a better appreciation of those specific characteristics which are most essential and desirable in them. Teachers who still think of the "new-type" test as "only fact-finding" in character are hardly likely to be properly motivated, no matter what the true character of the tests provided.

It is equally if not more important that the *limitations* of present measuring instruments be more adequately recognized. Even the best of the tests now being provided fall far short of measuring all of the desirable outcomes of instruction in any field of subject matter. In general, they have thus far been concerned only with the rational and intellectual achievements of the pupils and have measured only certain aspects of this type of achievement. This in itself, of course, is no condemnation of the tests — the real danger lies only in the failure to recognize that these limitations exist. In the opinion of the writer, it is one of the major responsibilities of those conducting coöperative testing programs to educate test users to a more adequate recognition of these limitations. School administrators particularly should be warned that the measures secured from the tests constitute only a *partial* basis for evaluating the total effectiveness of instruction, not only because they fail entirely to measure some of the outcomes of instruction, but because pupil achievement, as measured by these tests, is influenced by many factors entirely beyond the control of the individual teacher.

The influence upon instruction of any specific testing program in any school situation primarily depends, then, upon the nature and quality of the examinations provided and upon the intelligence with which these examinations are used. In conclusion, therefore, it may be noted again that, because of the variations in existing programs in these respects, no safe generalizations can be made concerning them. The writer is convinced that in many of the programs now being conducted the consequences are less often desirable than otherwise, and that there is serious need for a more critical attitude toward organizations of this type on the part of school people in general. It is equally true, how-

ever, that certain of the programs are on the whole providing a very valuable service and exercising a wholesome influence in the participating schools. If continued efforts toward improvement are aided by constructive and discriminating criticisms, the coöperative testing program should eventually become a permanent and valuable addition to our whole system of public education.

Summary and Reactions

Clifford Woody
University of Michigan

The writer, at whose suggestion this symposium originated, feels constrained to review briefly some of the generalizations offered by those contributing and to inject a few comments setting forth his own idea on various of the issues involved. Any attempt to reduce the statements made by any group of high-powered individualists and specialists, such as represented by the contributions to this symposium, is a hazardous undertaking and may result in inadequate statements of the exact positions set forth by the individual authors. The writer, if he does make apparent misstatements, will offer two alibis: (1) that the necessary briefness of the reports precluded the authors from setting forth vividly and in totality their positions concerning the various issues germane to the problem; (2) that the writer's mind-set and convictions concerning measurement and its influence on instruction are such as to incapacitate him from interpreting the contents of the various papers presented according to the intent of the authors. Since every reviewer is the victim of these two conditions, each reader is urged to study the articles, to make his own generalizations and draw his own conclusions.

Since this article is intended to be in the nature of a summary, the various generalizations made will be set forth in enumerated fashion.

1. *All authors recognize the positive potential contribution of measurement as a means for the improvement of instruction.* The authors may have defined measurement in slightly different terms according to the various interests presented, but they universally pointed out that if the purposes are properly conceived, if proper techniques of measurement are employed, and if proper uses are made of the results, beneficial results will eventually follow in the use of measurement in the instructional scheme. It may be that this apparent agreement of the possible beneficial effects of measurement on instruction is more apparent than

real. It may be that the authors accept the possibility but doubt the probability. The saviors in the foregoing statements are rightly conceived "proper techniques" and "proper uses." Each author, no doubt, reserves the right to define these terms in his own way, yet the reviewer is of the opinion that there will be no wide differences in their definitions. Thus it would seem, in the opinion of this group, that measurement represents a legitimate process of studying instructional processes and offers opportunity for positive contributions on the improvement of instruction. To the writer, who thinks of measurement as a movement for reducing subjectivity to a minimum and increasing objectivity to a maximum in the process of making a judgment, no other position than the one taken by this group is justifiable.

2. *All seem to agree that the objections to measurement arise from faulty techniques of test construction, faulty selection of content, and faulty interpretation of results.* The authors, as stated under generalization "1," all accept the idea of measurement but they are all critical of some of the results which have accrued from efforts at measurement. The authors point out that some tests do not measure what they seem to measure, i. e., some do not give valid measures, or yield reliable measures, yet important decisions are based upon the results from the administration of such tests. Pupils have been reclassified, courses of study may have been modified, or the success or failure of a teacher may have been determined upon the basis of such test results. No doubt much abuse has arisen from this misuse of test results. Much of the abuse is probably contingent upon the apparent magic in the new instruments of measurement. To the novitiate and the uninformed, measurement is a sort of a patent medicine. It is something of a cure-all for all sorts of educational diseases regardless of the causes or the conditions under which the diseases develop. Without prolongation of this analogy, it may be added that some of the severest critics of measurement charge that test enthusiasts too often measure to see at what level the educational disease may be without exerting very much energy to remove the causes conditioning the disease. The authors almost universally recognize the abuses accompanying the measurement movement and argue for the common-sense use of tests in the hands of especially trained specialists who in turn must exert maximum effort to enlighten all persons, connected directly or indirectly with tests or test results, as to their true meaning and interpretation.

3. *The authors also seem to agree that the abuses and misuses of measurement present a challenge for improvement in the field rather than abandonment of measurement.* The articles adversely criticizing the testing movement are timely and valuable but they may leave the impression that there is so much badness in the movement that further activities involving it should be taboo. The emphasis on the abuses is so great that these critics develop in the mind of the reader a condemnation of the whole movement rather than of the abuses coincident with the movement. These critics may unwillingly be killing the educational goose which promises to lay the scientific eggs necessary for the development of a science and a profession of teaching. On the other hand protagonists of the movement in their publications or verbal pronouncements are so inclined to present the contributions of the values accruing from the utilization of measurement and to refrain from mentioning the limitations in the instruments of measurement or in the uses thereof that the reader or hearer becomes ultra-enthusiastic in his acceptance of the movement and blind to its shortcomings. The contributors to this symposium have avoided the extreme position of the destructive critic with his over-emphasis on abuses and limitations and that of the protagonist with his ultra-enthusiasm on contributions and possibilities. These contributors attest belief in measurement *per se*, recognize clearly the deficiencies, defects and dangers, and then accept these limitations as the challengers of the movement. They call for the scrapping of old and inadequate tests and for the building of new and more adequate models. They condemn faulty practices in the administration of the tests and in the interpretation of the results, but at the same time point out the ways and means to avoid the malpractices. The contributors in this symposium have thus presented the problems of the measurement movement in such way as to present a challenge to its healthy development.

4. *Most of the participants in this symposium pointed out the need for development of instruments of measurement for subtle and dynamic qualities such as initiative, coöperation, methods of work, methods of thinking, appreciations, attitudes or emotional effects.* Measures of these types of educational products are essential in order to give consideration to the totality rather than to the limited portion of the effects resulting from an educative experience. There was repeated reference to this need for additional types of measurements, although the contributors recognized the development of these desirable types of instruments of meas-

urement as fond hopes rather than present realizations. One of the contributors states that all values to be emphasized in teaching should be incorporated in the tests designed for the effectiveness of such teaching. This contributor implies that if teaching values can be differentiated and defined, these same values can be included in test construction. He contends that in the construction of achievement tests for state-wide appreciation, for which he is responsible, he is dominated by the values emphasized in the existing courses of study and in the teaching practices of the state. Another contributor, while holding the view that tests for this so-called intangible values are needed and that if such outcomes to be emphasized in teaching can be so clearly delineated that they can be measured, points out that the difficulty hindering the development of such instruments of measurement is the fogginess surrounding the values to be emphasized in the teaching of the various subjects. This contributor states this point in a rather terse way when he says, "The failure of examinations (he was speaking of achievement examinations rather than measurement in general) to represent higher objectives is perhaps due to *haziness* of objectives when translated into behavior expected of the student . . . If the objective can be stated in teachable elements, student achievement can be measured by sampling those same elements." Such statements indicate a need for wider testing and a promise that such testing will result when the instructional atmosphere surrounding the different subjects clears.

The writer has long advocated emphasizing in the instruments of measurement all the values which should be emphasized in teaching. He has long argued that the first step in test construction is a determination of the values which are worthy of being measured. This means a most critical analysis of the values to be emphasized in the teaching of any given subject. It seems that unless such values are emphasized in the instruments of measurement they will not be emphasized in teaching. He is convinced that instruction tends to follow the lines of emphasis indicated in its appraisal. The testimony of the contributors to this symposium seems to corroborate this contention. Some praise the measurement movement because consideration of the values and items stressed in instruments of measurement have had beneficial effects in causing teachers to emphasize higher values in this instruction; others have condemned the movement because the effect of the values and content stressed tended to cause emphasis on teaching lower type

mental values and antiquated items of subject-matter. Such statements surely attest to the influence of measurement on instruction and suggest the necessity of making test values coincident with instructional values.

While the writer accepts the necessity for development of instruments of measurement stressing the same values as should be stressed in instruction, he realizes the difficulty involved in such efforts. Any test maker who is brazen enough to attempt setting forth these values in such clear outlines that "he who runs may read" is almost taking his educational life in his hands. There is such disagreement on the values and content to be stressed that even national committees composed of the nation's outstanding specialists in the fields concerned and endowed with large sums of money for investigational purposes have hesitated to set forth clearly programs of values to be emphasized in the various fields concerned. Complacency exists so long as the values are enmeshed in this nebulous matrix, but the test maker if he emphasizes a broad program of teaching values in his instruments of measurement, must of necessity outline in vivid detail the type of teaching situations designed to stress these given values. He thereby calls upon himself the criticisms of hordes who either disagree in the statement of the desired values or in the teaching situations selected. Such a situation may aid in clarifying the objectives underlying the teaching of the given subject, but it certainly makes the task confronted by the test maker a difficult one. The test maker must come into the open and by so doing make himself the target for various types of educational sharp shooters. The test maker must either limit his efforts to those simple values upon which there is wide-spread agreement or venture at his peril into the values of no-man's land. The contributors to this symposium seem to urge the need for following the latter course. They do caution, however, that the limitations encountered in following such a course should be clearly delineated.

5. *The contributors implied, if they did not implicitly state, that the various types of measurement must be evaluated in terms of the purposes for which they are designed.* There are numerous types of testing programs and various types of tests. These testing programs and types of tests are designed to serve different purposes. In this symposium one of the participants discussed measurement from the standpoint of a given state-wide every-pupil examination in which the dominant purpose is the determination of the degree of mastery of the

subject-matter and disciplines emphasized in the adopted materials of instruction; another presented the view of a non-instructional appraisal agency given the responsibility for evaluating achievement within an institution of higher learning; another presented the tests and testing programs as instruments designed to be given and used by teachers themselves; still another wrote from the point of view of one who uses tests as instruments for discovering as broad and as accurate information as possible concerning an individual or a pupil. Still other points of view were expressed, but a sufficient number have been enumerated to suggest that testing programs and tests serve vastly different purposes. Since tests are designed for different uses and are conditioned by the circumstances prevailing at the time the program or tests are contemplated, the publication of the tests and the results makes possible the use of the instruments in situations not comparable to those for which the instruments were designed. This more or less indiscriminate use of tests without consideration of the functions which they are best fitted to serve is responsible for much of the criticism leveled at the testing movement. Such indiscriminate use, usually accompanied by wild speculation as to the meaning of the tests, is often referred to by the participants in this symposium as the unintelligent use of tests.

Failure to take into consideration the purposes to be served often leads to confusion concerning the use of tests and testing programs. Some tests, administered by agencies outside the instructional staff, i. e., by a group of specialists selected for making certain appraisals within the units under consideration or by a group of specialists within a clinic, may be designed to determine peculiar strengths and weaknesses unrelated to values emphasized in the instruction within the units concerned. The aim of such tests may be to determine levels of achievement primarily for other than instructional purposes, although instruction may be indirectly affected. Such tests may be given for purposes of classification of students, or for determining variation in the achievement of groups of pupils, or for determining the level of achievement on various items foreign to the materials of instruction usually stressed. This type of measurement may be undertaken when the major interest lies in the measurement itself, divorced from any direct effect on instruction or instructional processes. Such type of measurement has ample justification and its function should be recognized. However this type should not be confused with that designed primarily for the improvement of instruction. The plan of administra-

tion of the latter type and the utilization of the results therefrom is entirely different from that of the former. The latter techniques are designed primarily for the purposes of influencing instruction; the former, for non-instructional purposes.

Much of the unfavorable criticism of the influence of tests arises from the failure to make differentiations in the functions of the tests. The reader will note that the contributors to this symposium are usually very careful to confine their remarks to a specific kind of test or testing program. Their generalizations are not intended as general applications. When the various statements are considered in the light of the situations set up, it is ventured that most of the conclusions set forth will find acceptance by the different authors.

6. *The authors seem to imply that values emphasized in the tests must change from time to time and must be consistent with the essential values of the educational process.* It has been charged that measurement concerns itself with *what is* and not *what ought to be*. This, it contended, makes for *status quo* rather than for future progress and development. One author stresses the fact that measurement instead of contributing to *status quo* has provided the stimulus for redefining aims of instruction, improved selection of subject-matter, modification of methods of teaching, and even for refinements in the tests themselves. He avers that the use of the tests has stimulated rather than retarded educational progress. Other authors emphasize the necessity of keeping the tests abreast with educational progress and suggest that as soon as educational objectives are clarified and distinctly set forth, appropriate instruments of measurement will be devised to determine the extent to which the objectives are being realized.

Present examination of many instruments of measurement developed in the past certainly suggest some out-worn values. However, it must be recalled that many of these devices were in the infancy of the measurement movement when its techniques and philosophy were in the formative stages. Many of these early devices, useful as they were at the time of their development, are imperfect and should be discarded in favor of improved models just as old automobiles are exchanged for new ones. But in criticizing these old tests, one should keep in mind that probably the instrument itself may not have been so much at fault as was the prevailing system of educational values which gave birth to the idea of the device. It is probable that the measuring device was but a reflection of the then acceptable philosophy of educa-

tion. Certainly the prevailing systems of educational values were utilized in the interpretation and evaluation of the results obtained. This system of values, no doubt, was not universally accepted and was not sanctioned by certain minorities or blocs assuming various roles of leadership, but it was surely prevalent among the rank and file of educational workers. Thus the instruments of measurement developed in the past reflect the educational thinking of the age in which they were developed. As social conditions change and new educational values evolve, new instruments of measurement reflecting the newly evolved values will surely follow. Condemning the instruments of measurements as forces making for *status quo* is merely saying that the philosophy of education providing the media in which those instruments were developed is unsuited for present-day social and educational needs. Old philosophies of education and old test values must be discarded and new philosophies and new tests must incorporate the values emerging from an evolving society. If the writer properly interprets the import of the accounts in this symposium, the contributors accept the thesis that the values of measurement must be ever-changing in keeping with the changes of an evolving philosophy of education.

While it seems to the writer that the six enumerated generalizations as inferred from the content of the articles are stated in a positive and optimistic form, it should be added that the contributors are keenly aware of the defects of the movement. One contributor fears that cooperating state-wide testing programs will move the supervisory leadership to an outside agency; one condemns outright the "determining of the old tyrant *norm*" and states that the "administrative tests of the usual type should be discontinued or so managed as not to affect teaching and teacher morale adversely;" still another warns that "the character of the tests employed in measuring the results of our teaching influences the instruction and that the increasing use of objective tests is tending to nullify our attempts to change the curriculum." The writer does not interpret this sampling of statements as a wholesale condemnation of the testing movement as having a pernicious effect on instruction, but rather as a warning against the unwise selection of tests and test procedures and as a challenge for the production of better instruments of measurement, better techniques of test construction, and better methods of using test results.

EDITORIALS

MEASUREMENT AND EVALUATION

THE symposium on the foregoing pages of this issue on the "influence of measurement on instruction" is timely. The availability of three or four thousand standardized instruments of measurement and countless numbers of the so-called informal or new-type tests, the establishment of research bureaus or other agencies given responsibilities for the direction of testing programs of various proportions, the pressures from outside agencies for wholesale participation in sectional or national research projects, and the solicitation of publishing houses for the purchase of test wares should force to the front the question of the place of measurement in the scheme of instruction. Since the different contributors presented in acceptable fashion the advantages and disadvantages to be gained from the utilization of tests and testing programs there is no need for a restatement. However, from the point of view of the writer the factor determining whether or not measurement will guide to beneficial or detrimental effects depends upon the educational values or philosophy of education of those responsible for the direction of the program and the interpretation of the results.

The data obtained from tests and testing programs are one thing; the determination of the meaning and significance of these data is another. The scales may show that an individual weighs one-hundred-eighty pounds. Assuming that the scales weigh accurately, the weight of one-hundred-eighty pounds is a fact. However, the meaning and significance of fact for the individual is another matter entirely outside of this fact itself. The man may be of normal weight, underweight or overweight, depending upon a number of conditions, circumstances and upon the best available information and beliefs concerning health conditions. What the individual does about the fact depends in last analysis upon his particular beliefs and convictions concerning life values. In other words, his behavior in response to the information about his weight is determined by his philosophy of life.

Likewise the accepted philosophy of education functions in the determination of the meaning and significance of tests and data obtained

therefrom. One's philosophy of education should be and is manifested in the construction or selection of the instruments of measurement. It certainly is paramount in the interpretation of the data obtained from the use of the tests. Assume an administrator wishes to determine the level of achievement in addition, and that after careful examination of the available tests he chooses the one best fitted to measure the skill in question. It would be presumptuous to infer that he chose the test without regard to his belief concerning fundamental educational values. Assume also that upon the administration of the test he found the pupils under consideration solved eight out of twelve problems correctly but that their percentage of accuracy was only eighty percent. These findings are facts, but whether or not they are significant must be interpreted in terms of educational principles and values outside of the facts themselves. One's philosophy of education may say even if the level of achievement is on a par with that achieved throughout the country, that the condition portrayed in the particular class is satisfactory or unsatisfactory. One's beliefs and convictions, even though based to as great an extent as possible upon scientifically determined conclusions will dictate whether or not anything shall be done about the situation portrayed.

If the above contentions are sound, it should be clear that measurement results in data descriptive of existing situations; and that appraisal of the significance of the data results from the functioning of prevailing educational and life values. The failure to make these distinctions in descriptive data and appraisal thereof results in much confusion in the thinking about educational measurement. It is the conviction of the writer that the existence of a more comprehensive and clear-cut philosophy of education will prevent much of the misuse and abuse in practices involving educational measurement. Something may be and is wrong with measurement, but in last analysis the fault lies with a philosophy of education which will allow such malpractices to exist.

CLIFFORD WOODY

EDUCATIONAL DIAGNOSIS*

The purpose of this editorial is to call attention to the Thirty-fourth Yearbook of the National Society for the Study of Education, and its splendid contribution to the scientific study of classroom instruction.

*Brueckner, Leo J., and Others. "Educational Diagnosis," *Thirty-Fourth Yearbook* of the National Society for the Study of Education (Bloomington, Illinois: Public School Publishing Company, 1935), x + 563 pp.

The Yearbook is divided into five sections: (1) Factors associated with learning difficulties, with chapters on the physical factors, the intellectual factors, the pedagogical factors, the emotional and social factors, and the environmental factors associated with or conditioning learning; (2) principles and techniques of educational diagnosis with chapters on the characteristics of satisfactory diagnosis, elements of diagnosis, technique of diagnosis, the place of the psychological laboratory in educational diagnosis, maturation as a factor in diagnosis, and the principles of developmental and remedial instruction; (3) diagnosis and remedial instruction related to the various fields of instruction, with chapters on the diagnosis and treatment of disabilities in reading, English, arithmetic, social studies and science; (4) diagnosis and remedial treatment related to other objectives of instruction, with chapters on diagnosis in health education, behavior disorders, speech, vocational interests, abilities, and aptitude, musical talent, art, leisure-time activities, and creativeness; and (5) administrative aspects of diagnosis and remedial instruction, with one chapter on the administration of a program of diagnosis and remedial instruction.

Besides furnishing the reader with a generally helpful discussion of educational diagnosis, the committee responsible for the preparation of this Yearbook has made particularly worth-while contributions to the practice of diagnosis: (1) by indicating the relation of diagnosis to the objectives of education (this is important because of the tendency in some quarters to associate educational diagnosis with only those deficiencies experienced by pupils in the mastery of the school subjects); (2) by distinguishing between measurement and diagnosis (measurement supplies data for diagnostic reasoning; helpful diagnoses are to be made not through the mechanical application of tests but through carefully controlled thinking about carefully collected data); (3) by stressing the importance of the validity, reliability, and objectivity of diagnosis (definite suggestions are offered for the application of these and other criteria to educational diagnosis); and (4) by indicating in some detail the application of the diagnostic technique to practical classroom situations (a number of chapters are devoted to the discussion of the diagnostic and remedial treatment of deficiencies experienced by pupils in classroom situations.)

Aside from certain matters of detail as, for example, the committee's failure to supply the reader with any very definite notion of the nature of diagnostic thinking, the only major omission sensed by the writer in

reading this report is the absence of any discussion of the relationship between well planned first instruction and diagnosis. And involved in this practical relationship is the more fundamental theoretical relationship represented by the two well known present day schools of educational psychology: (1) that which is concerned with the development of the general law, principles, and conditions for efficiency in learning; and (2) that which is concerned with the direction of the mental efforts of children in specific learning situations. While the two approaches are not incompatible, the writer feels, in the absence of any thorough-going discussion of the relatedness of these two rather fundamentally different approaches to learning, that many persons may come to a decision about their relative merits and that of well-planned first instruction vs. diagnosis, not from any consideration of the real issues involved but from other considerations, general impressions, analogies, or what not. For those who can bridge this gap, the materials furnished by this Yearbook will be found educationally sound and workable.

Again, the writer wishes to commend to the readers of the *Journal of Educational Research* this very realistic and scientific approach to the problems of classroom instruction.

A. S. Barr.

REVIEWS

Address all communications relative to book reviews to Editor Kai Jensen, University of Wisconsin, Madison, Wisconsin.

DUNLAP, KNIGHT. *Civilized Life: The Principles and Applications of Social Psychology* (Baltimore: The Williams and Wilkins Company, 1934), pp. IX + 374.

This volume is a revision and extension of Dunlap's Social Psychology published in 1925 and includes new chapters on Desire, Race and Civilization, and the Child. Dunlap's method is to discuss both the function and the explanation of social phenomena in terms of human desires. These, he suggests, may be adequately classified under nine headings, (1) alimentary desire, (2) excretory desire, (3) protection desire, (4) activity desire, (5) desire for rest and relaxation, (6) amorous or erotic desire, (7) parental or philopedic desire, (8) desire of preeminence, and (9) desire of conformity. Every actual desire is a complex process involving a particular type of thinking, and something more. This something more he calls the "appet." In some desires the appet can be identified as an organic condition. "Desire for food and desire for drink are conditioned primarily by hunger and thirst." In most desires the appet is obscure.

With this general beginning human institutions, including marriage and the family, religion, and government are explained and the values of particular forms assessed in terms of their contribution to the satisfaction of desire. The "principles of social function," including communication, social organization, morals and the forms of social control are discussed in the same terms.

This method has, of course, a great advantage over the method of McDougall which attempted a description and explanation of human social behavior in terms of a list of instincts. Dunlap's desires are much more manageable than instinctive tendencies because they are admittedly the results of learning and in them varied forms are to be expected.

But the method leaves a great deal to be desired. No criteria are offered by which it can be judged whether or not any particular action is an instance of one or the other of the nine desires. Here is, for instance, a man building a house. Only through a knowledge of his thoughts could it be told whether he is driven by an appet centered in hunger spasms, or whether protection is the motive, or whether he is suffering from a mere desire for activity, or earning a chance for rest and relaxation, or a chance to marry a wife, or to have a family, or to be a preeminent carpenter, or all or none of these. Dunlap scorns the behavioristic method even in the field of social psychology where (unless we are credulous of stories of telepathy) it might well be urged that the individual affects others only through his more gross movements.

The outcome of this treatment of human social behavior in terms of desire is a record of Dunlap's opinions on a wide variety of human affairs. These

opinions are very readable and interesting, but they remain, for the most part, opinions.

E. R. GUTHRIE
University of Washington

JERSILD, ARTHUR T. *Child Psychology* (New York: Prentice-Hall Inc., 1933), pp. XIII + 462.

The plan of Jersild's new text in child psychology is the usual one of considering the capabilities of the child at birth, and later taking up the development of the child in separate behavior categories, most of which have the same titles as chapters in elementary texts. Thus the first chapter describes the newborn infant, omitting any discussion of prenatal development. Most of this chapter is drawn from work directed by the late Professor Weiss (Pratt, Nelson, and Sun; Irwin; Jensen; and Marquis. A second chapter discusses the development of motor coordination, locomotion, and handedness, and includes part of Gesell's developmental norms. The remaining chapters, with the exception of the last one, describe in greater detail the development of the following aspects of the child's behavior: language, emotions, social behavior, learning, understanding, mental ability, and personality.

The final chapter, "Some Aspects of Applied Child Psychology," discusses problems of child training both generally and specifically. The general discussion includes statements of the need of discipline and consistency in the treatment of children; and comments on the role of maturation in development; and a plea to coördinate disciplinary training with the child's interests. The specific discussion gives suggestions on training habits of sleep, eating, elimination, and so on.

The emphasis in all chapters is on the child of pre-school age.

Throughout the entire text the importance of maturation is stressed as an explanation of childhood improvement. In several instances experimental records are cited which show that a control group gained as much or nearly as much in a given ability as an experimental group. For learning tasks not credited to maturation the concept of reintegration is used.

There may well be some criticism of the author's selection of materials. Studies from the Child Development Institute are very frequently quoted. On the other hand there are some surprising omissions. Although the chapter entitled "The Growth of Understanding" is 65 pages long, there is no reference to the important work of Piaget, nor is he quoted in the bibliography of references on that subject.

There is a bibliography at the end of each chapter. Citation of the authors referred to within a chapter is done by name only. Both subject and name indexes are included.

In difficulty the book is suited to use in a beginning course on child psychology. The students in such a course will probably react favorably to the book because of the relatively simple and clear style of writing and the lack of difficult and detailed experimental results. From the standpoint of the professional psychologist Jersild's *Child Psychology* is not as important a contribution to the field as some other recent works.

DAEL L. WOLFLE
University of Mississippi

CLEMENSON, JESSIE WILLIAMS. *Study Outlines in Physics Construction and Experimental Evaluation.* Contributions to Education, No. 553 (New York: Bureau of Publications, Teachers College, Columbia University, 1933), pp. IX + 154.

A study dealing with the oft recurring question:—"What's wrong with secondary school physics?" with a suggestion in teaching technique aimed to correct some of the outstanding difficulties in the accomplishment of desired ends in the teaching of the subject, is the thesis of a recent publication, entitled "Study Outlines in Physics."

The report of the study is made in two parts. Part one concerns the construction of the study outlines in Physics. The author has made and reported on a survey of several current philosophies of education, the aims and objectives of physics teaching, and the psychology of the learning process. The study outlines were constructed to incorporate these philosophies and to make the materials of physics available to teachers and pupils in a form that would provide for the operations of the laws of learning. To this end the subject was grouped about eight units; these covering the field of high school physics. Each unit was analyzed into its main ideas and concepts. Questions were then formulated on the concept for development of the ideas. The question is in problem form and is worded and arranged to develop interests, attitudes, and habits as well as concepts, knowledges and skills.

Part two is a report of the experimental procedure adapted to evaluate the study outlines. Nine large high schools of Los Angeles, California, coöperated in the experiment in an effort to answer such questions as: Are such outlines usable by the average teacher in the average class room situation? Do the outlines affect the pupil's mastery of facts and concepts? What effect do they have upon study habits, and upon his attitude toward problem solving and his interest in his environment? Insofar as the data given are concerned, the study outlines, developed by the author, and used in these schools have very positive values and are a step in the direction of attainment of some very desirable goals in physics teaching.

An appendix is provided which is of value to a teacher who plans to make use of the suggestions of the study. These materials, here presented, will enable him to keep in mind the objectives of high school physics and will give many suggestions as to how the materials of the units may be made to contribute effectively to these objectives. Sample outlines are given with suggestions for their construction and use.

The study is convincing in its conclusions. Progressive teachers of physics will find in this report many profitable suggestions for improving teaching technique.

VERGIL C. LOHR
University of Chicago

WILSON, HOWARD E. *The Fusion of Social Studies in Junior High Schools.* Harvard Studies in Education, XXI (Cambridge: Harvard University Press, 1933), 211 pp.

During the last twelve or fifteen years the literature of education has contained many articles advocating or opposing the fusion, combination, or synthesis of the social studies so as to make them more lifelike and functional. The articles referred to, with a few exceptions, have been partisan in character, expressing the views of advocates or opponents rather than presenting a disinterested, objective evaluation of the merits and weaknesses in the fusion proposal. Now, for the first time, a volume is available which in large measure provides a comprehensive survey and scholarly criticism of this phase of curricular experimentation. It in no way detracts from the

value of the monograph to point out that some of the ideas it contains have appeared in earlier publications by the author.

Dr. Wilson states that the purpose of his study is, first, to determine what the theory of fusion is, and, second, to analyze the theory so as to estimate its value. To accomplish his initial purpose he first analyzes ten outstanding fusion programs: namely, those in the Laboratory School of the University of Missouri (now abandoned); the Marshall-Judd program (not completed); and the programs in the Junior High Schools of Fond du Lac, Wisconsin (later revised); Beaumont, Texas; San Jose, California; Cleveland, Ohio; Horace Mann School, Columbia University; San Antonio, Texas; Long Beach, California; and the Lincoln School, New York City. He then analyzes the writings of the advocates of fusion (Rugg, Ball, Seymour, Hatch) and fusion critics so as to discover what the fusion theory is. By the foregoing method of investigation he derives the following tenets or principles of the fusionists (p. 61):

1. Only such material as has direct value in developing in pupils intelligent understandings and tolerant, coöperative appreciations fitting them to engage in the activities of the life of the time shall be taught.
2. Selected subject matter in the social sciences must be organized in units of experience, psychologically appealing and learnable, and corresponding as closely as possible to life situations.
3. Traditional subject boundaries shall be ignored in the construction of the social-science curriculum; subject fields not only fail to achieve purposes of education but interfere with the selection and organization of a curriculum which will achieve these purposes. The current problem rather than the subject is the heart of a functional unit.

The rest of the monograph is devoted to an evaluation of the theory of fusion as set forth in the foregoing principles by an examination of the historical conditions under which fusion evolved and against which it revolts, the author analyzes fusion from the viewpoint of the social sciences and the "boundaries" laid down for them by the subject specialists. This analysis shows convincingly that the merits of fusion must be weighed, first, against the traditional course of study in the social sciences and, second, against the progressive "subject" courses; or, if the traditional be eliminated, the contest must lie "between fusion in 1930 and non-fusion in 1930, not between fusion in 1930 and non-fusion in 1900" (p. 105). An examination of recent subject courses and textbooks indicates that the fusionist assumption that subject courses include much nonfunctional material and exclude much useful material is not warranted. So far as the fusionist method of organizing material is concerned, an examination of fusion units and subject units found in progressive schools leads to the conclusion that, in general, the subject theory of organization seems superior to the fusion theory, although—the author adds—"it must be recognized that the elements on which the superiority is based are primarily the result of historical development, possibly of historical accident" (p. 179).

The final conclusions of the author are as follows: (1) "The theory on which subject teaching is based would seem to offer greater educational opportunities than does the fusion theory, although the differences . . . are not sufficient to

rule either theory out of consideration'' (p. 185); (2) ''there is no guarantee in either the fusion theory or the subject theory that courses constructed according to the dictates of either will thereby achieve their greatest potentialities'' (p. 186); (3) ''subject courses, on the average at present, may more nearly approximate their ideal than fusion courses may because practical considerations of teacher-preparation and general curriculum scheduling are favorable to subject teaching'' (p. 187); (4) ''the fusion theory is highly serviceable to social-science teaching in that it emphasizes educational and cultural concepts which subject specialists frequently forget and, in its own right, is useful as an occasional variant from the subject approach'' (p. 190).

Finally, it is not too much to say that in this study Dr. Wilson has made a valuable contribution to educational literature and, in particular, to the literature of social-science curriculum construction.

HOWARD C. HILL
University of Chicago

KIMBALL, ELSA PEVERLY. *Sociology and Education: An Analysis of the Theories of Spencer and Ward* (New York: Columbia University Press, 1932), 323 pp.

It was a happy moment when the author conceived the idea of contrasting the educational theories of Spencer and Ward. Not only were they the major representatives of the two conflicting social philosophies of the nineteenth century; their opposite solutions of the problems of social happiness still remain as *laissez faire* versus social telesis separate our people into two uncompromising groups. Spencer and Ward were alike in their insatiable love of the seeking knowledge of human experience, their abandon in social interpretation, their confidence in the finality of their doctrines, and the seriousness with which they regarded it their duty to spread their teaching. In the fundamentals of their systems they could not have been farther apart. In their attitudes toward education they came closest together. Ward[1] realized this, as appears in the address he prepared but did not deliver at the complimentary banquet given Spencer just before he left this country to return to England.

Although it is the purpose of this book to contrast the educational theories of these two representative thinkers, most readers will value it chiefly for its discussion of the background of Spencer and of Ward and the author's analysis of the influences that led Ward and Spencer to educational ideas that had much in common in spite of the philosophic differences of the two men.

It is difficult to see how anyone of critical disposition, close to the practices of public schools in the United States, can avoid feeling depressed when he realizes how far our education still is from the ideals of either Spencer and Ward. This comes out when one considers, for example, the plea of Spencer for the education of parents and the argument of Ward for education for marriage and for parenthood even more ahead of the common thinking of his time and of ours. Ward,[2] however, had his eye here as always on the distant but better future, as appears in his closing sentences in his *Pure Sociology*, the most significant contribution yet made in his science by American thinkers:

''The action of society in inaugurating

[1] Ward, Lester Frank. *Glimpses of the Cosmos*. Edited by E. P. Cape (New York and London: G. P. Putnam's Sons, 1913-1918), pp. 114-117.

[2] Ward, Lester Frank. *Pure Sociology*, New York: The Macmillan Company, 1903.

and carrying on a great educational system, however defective we may consider that system to be, is undoubtedly the most promising form thus far taken by collective achievement. It means much even now, but for the future it means nothing less than the complete social appropriation of that individual achievement which has civilized the world. It is the crowning act in the long list of acts . . . that constitute the socialization of achievement.''

ERNEST R. GROVES
University of North Carolina

SYMONDS, PERCIVAL M. *Mental Hygiene of the School Child* (New York: Macmillan Company, 1934), pp. XI + 321.

The maintenance of good mental health.—For many years the public has become increasingly aware of the tremendous economic and social waste due to the large number of individuals who break down mentally. Institutions for the mentally abnormal and for delinquents and criminals are rapidly taxing the resources of our cities and states. Within the past twenty years students of human behavior have recognized that many personality deviations originate in childhood. In consequence clinics have been formed in a large number of cities, psychologists have been appointed in many school systems and educational counsellors have been increased in number. And yet little progress has been made principally because these measures have functioned essentially as therapeutic rather than preventative.

With the widening scope of education the school has assumed many additional functions. Mental hygiene has become one of the problems with which the classroom teacher is intimately concerned. A large number of books have been published with the purpose of informing teachers of the problems of mental hygiene and of classroom practices. Of the recent books that by Professor Symonds of Columbia University is directed essentially to the classroom teacher rather than to the specialist in the field.

The author states in his preface that the book is unusual because of its emphasis on the positive aspects of mental hygiene. Nevertheless, many pages are devoted to an evaluation of negative or undesirable behavior. Furthermore, the positive mental hygiene mentioned in the preface is not always convincing. Such topic headings as ''A child should learn to have a glow of satisfaction in doing things well'' (page 84), and ''A child should learn to be sensitive to the good, the true, and the beautiful'' (page 89), do not properly belong in a book on mental hygiene. On page 90 the author states, '' . . . a child learns to be sensitive to the beautiful by having pointed out the beauties of a warbler, a colonial house, a prairie sunset, or a minuet.'' Granted that a sense of beauty can be instilled in children in such ways, there is no reason to believe that it would aid in the maintenance of normal mental health.

Beginning with Chapter X a number of important topics are discussed and constructive suggestions are given. Chapter X contains excellent suggestions on the problem of sex information and sex adjustment. Chapter XVII on remedial work likewise contains interesting and useful information and should be read by every teacher.

The most convincing discussions are contained in Chapters XVII to XX. Chapter XX contains a presentation of six case studies rather well done.

Despite occasional superficiality and some dogmatic assertions, the book is well

worth reading. It should be of special value to teachers whose training has not included advanced courses in psychology. The book will also serve to make teachers who have done little reading in the field of mental hygiene acquainted with some of the problems of the field and with some classroom practices. It will probably be of greater use to parents than to trained teachers. Teachers can be of genuine service to parents by recommending this book to them.

MANDEL SHERMAN
University of Chicago

RESEARCH ABSTRACTS and BIBLIOGRAPHIES

Address all communications relative to research abstracts and bibliographies to A. S. Barr, University of Wisconsin, Madison, Wisconsin.

BETTS, E. A., AND BONTRAGER, O. R. *Research Studies in Elementary School Language No. 1* (Iowa City, Iowa: University of Iowa, 1934), 61 pp.

This bulletin contains reports of two investigations: (1) An Evaluation of Certain Techniques for the Study of Oral Language, by E. A. Betts; and (2) An Experimental Appraisal of Pupil Control of Certain Punctuation Items, by O. R. Bontrager.

CRAWFORD, J. R., AND KIRKPATRICK, J. E. *Research Studies in Educational Measurements I* (Iowa City, Iowa: University of Iowa, 1934), 64 pp.

This bulletin contains reports of two investigations: (1) Age and Progress Factors in Test Norms, by John Raymond Crawford; and (2) The Motivating Effect of A Specific Type of Testing Program, by James Earl Kirkpatrick.

EARLE, F. M., AND KILGOUR, J. *A Vocational Guidance Research in Fife* (London: The National Institute of Industrial Psychology, 1935), 101 pp.

This bulletin furnishes a systematic account of the results and conclusions derived from them of the vocational guidance experiment conducted during the years 1928-1932 by the National Institute of Industrial Psychology in the county of Fife, Scotland. After describing the experimental setup data are offered of the urban and rural vocational conditions, the age at which vocational guidance should begin, and the minimal qualifications desirable in various kinds of work. Appendix.

FITZGERALD, JAMES A., AND OTHERS. *Doctoral Theses in Education I* (Iowa City, Iowa: University of Iowa, 1934), 202 pp.

This bulletin contains a report of six doctors theses as follows: (1) Letters Written outside the School by Children of the Fourth, Fifth and Sixth Grades: A Study of Vocabulary, Spelling Errors, and Situations, by James A. Fitzgerald; (2) An Experimental Study of Certain Techniques for Testing Word Meanings, by Victor H. Kelley; (3) Valid Testing and Diagnosis in the Mechanics of the Ninth Grade English Composition, by Raymond L. Powell; (4) Participation in College Athletics and Vocational Success, by Moses N. Thisted; (5) The Effect of Drill upon the Specific and the General Comprehension of Historical Content, by Thomas J. Tormey; and (6) The Concepts and Attitudes concerning Slander and Freedom of Speech, by Ida M. Yates.

IRWIN, ORVIS C., AND OTHERS. *Studies in Infant Behavior I* (Iowa City, Iowa: University of Iowa, 1934), 174 pp.

This bulletin contains the reports of five investigations as follows: (1) Differential Variations in the Amount of Ac-

tivity of Newborn Infants under Continuous Light and Sound Stimulation, by LaBerta A. Weiss; (2) The Effect of the Factors of Duration, Intensity, and Pitch of Sound Stimuli on the Responses of Newborn Infants, by Esther M. Stubbs; (3) Differential Variations in the Activity and Crying of the Newborn Infant under Different Intensities of Light: A Comparison of Observational with Polygraph Findings, by Orvis C. Irwin and LaBerta Weiss; (4) The Effect of Clothing on the General and Vocal Activity of the Newborn Infant, by Orvis C. Irwin and LaBerta A. Weiss; and (5) The Effect of Darkness on the Activity of Newborn Infants, by Orvis C. Irwin and LaBerta A. Weiss.

JACK, LOIS M., AND OTHERS. *Behavior of the Preschool Child* (Iowa City, Iowa: University of Iowa, 1934), 171 pp.

This bulletin contains the reports of five researches as follows: (1) An Experimental Study of Ascendant Behavior in Preschool Children, by Lois M. Jack; (2) A Study of the Development of Two- and Three-Year-Old Children with Respect to Play Activities; by Elizabeth M. Manwell and Ida Gaarder Mengert; (3) The Adequacy of Samples of Behavior Obtained during Short Observation Periods, by Esther Van Cleave Berne and Helen Garside Kelly; (4) An Experimental Investigation of Certain Factors Involved in the Preschool Child's Compliance with Commands, by LaBerta A. Weiss; and (5) A Study of the Behavior of Young Children in Anger, by Agnes Fairlie Ricketts.

MCCONNELL, T. R., HENRY, LYLE K., AND MORGAN, CLELLEN. *Studies in the Psychology of Learning, II*, Educational Psychology Series, Number 2 (Iowa City, Iowa: University of Iowa, 1934), 142 pp.

This bulletin presents three studies: (1) Discovery vs. Authoritative Identification in the Learning of Children, by T. R. McConnell; (2) The Role of Insight in the Analytic Thinking of Adolescents, by Lyle K. Henry; and (3) Characteristics of Problem-Solving Behavior of Adults, by Clellen Morgan.

MONROE, WALTER S., AND OTHERS. "Psychology and Methods in the High School and College," *Review of Educational Research*, IV (December, 1934), pp. 445-463.

The committee presents summaries of researches relating to the following fields: English language, Reading and Literature; Social Studies, Foreign Languages, Science, Mathematics, Commercial subjects, Industrial and Vocational Education, Fine Arts, Music, Health and Physical Education, Moral and Character Education, Home Economics, Education, Psychology, and Professional training of teachers. Bibliography.

NEWBURN, HARRY K., AND OTHERS. *Doctoral Theses in Education* II (Iowa City, Iowa: University of Iowa, 1934), 122 pp.

This bulletin contains a report of four doctoral theses as follows: (1) The Relative Effect of Two Methods of Vocabulary Drill on Achievement in American History by Harry K. Newburn; (2) The Difficulties of Making a High School Schedule of Recitations, by Roswell C. Puckett; (3) Objectives of Physical Education, by Edward Frank Voltmer; and (4) The Relative Effectiveness of Various Forms of Presentation of Dental Hygiene Reading Matter as Shown by Tests of Reading Comprehension, by Anna M. Wiecking.

OBRIEN, F. P. *The Public Junior College as a Community Asset.* A Survey Study for Pratt, Kansas. Kansas Studies in Education, Volume II, No. 4 (Lawrence, Kansas: University of Kansas, 1935), 41 pp.

After defining the nature of a Junior College the author discusses the need of a Junior College in Pratt, Kansas, the cost to the community, the service rendered to the community, with recommendations for the further development of the Junior College idea.

OSBORNE, DOROTHY M. *Emotional Instability;* A Case Study, Child Research Clinic of the Woods Schools, Langhorne, Pennsylvania, 1935.

This is the second of a series of case studies published by the Child Research Clinic of the Woods Schools, Langhorne, Pennsylvania, to further the study and help in the solution of the problems of the exceptional child.

Our Small Enrollment Rural Schools (Madison, Wisconsin: Wisconsin Teachers Association, 1935), 48 pp.

This bulletin contains data on the rural school enrollment for 1933-1934, small schools per counties 1929-1930, 1933-1934, number of large one-teacher schools per county, cost of maintaining rural schools in 14 counties, cost of transportation for schools in 14 counties, and materials relating to the rural school situations in specific counties.

EDUCATIONAL ADMINISTRATION AND SUPERVISION

BOLMEIR, E. C. "Recent Tendencies in Taxation for Public School Support," *The Elementary School Journal,* XXXV (February, 1935), pp. 415-422.

The author presents data on tax limitations per state, new types of taxes, and a discussion of state participation in school support.

MOORE, EOLINE, WALLACE. "Difficulties Recognized by Elementary Teachers," *Educational Administration and Supervision,* XXI (January, 1935), pp. 51-55.

The author presents from a study of the difficulties reported by 540 city school teachers and 544 rural school teachers a list of the difficulties most frequently experienced by elementary school teachers, promoting desirable habits, securing study aids, individualizing instruction, teaching reading, handling promotions, guiding study, teaching without texts, etc.

OLSON, OVE S. "Follow-up Practices in Teacher-training Colleges in Eight North Central States," *Educational Administration and Supervision,* XXI (January, 1935), pp. 45-50.

From a study of a questionnaire directed to one hundred teacher-training colleges in eight North Central states, the author finds that sixty-three of the sixty-four reporting schools answered the question, "Should teacher-training institutions have a system of follow-up?" in the affirmative.

SHANNON, J. R. "Academic Training of Secondary-School Principals." *Bulletin of the Department of Secondary School Principals of the N. E. A.,* No. 53 (December, 1934), pp. 7-12.

From a study of the training of some 2,000 principals the author concludes that a distinct advance in the training of secondary school principals has been made

over that reported by Eikenberry for 1922-1923. In that year the federal report showed 67.2 percent bachelors, 18.0 percent masters and 1.4 percent doctors. In this more recent report there are 44 percent bachelors, 42 percent masters and 1 percent doctors.

METHODS OF INSTRUCTION

BLAKE, MABELLE B., AND DEARBORN, WALTER F. "The Improvement of Reading Habits," *The Journal of Higher Education*, VI (February, 1935), pp. 83-88.

The authors present the result of an experimental study of an improvement program applied to good readers and poor readers at the college level.

MANZER, C. W. "The Effect of Knowledge of Output on Muscular Work," *The Journal of Experimental Psychology* (February, 1935), pp. 80-90.

From an experimental study of 68 men and 60 women college students divided into equal groups the author concludes from a study of the maximal muscular contractions as measured by the Smedley hand dynamometer that knowledge of output is followed by a prompt upward turn in the curve of muscular work.

PETERSON, H. A., AND OTHERS. "Some Measurements of the Effects of Reviews," *The Journal of Educational Psychology*, XXVI (January, 1935), pp. 65-72.

From a study of adult subjects taken from classes in elementary psychology in the Illinois State Normal University the authors conclude that the value of any given location of reviews depends on its nearness to learning on the one hand and to the time of reproduction on the other. These factors are compensatory in character.

SALISBURY, RACHEL. "Some Effects of Training in Outlining," *The English Journal*, XXIV (February, 1935), pp. 111-116.

From an experimental study involving 474 pupils in four high schools the author concludes that training such as was provided in this experiment will bring improvement in reading, an ability to reason, and achievement in content subjects to which either written or mental outlining may be applied.

CURRICULUM CONSTRUCTION

BRUMBAUGH, HARRY E., AND HUTSON, PERCIVAL W. "How High School Pupils Use the General Curriculum," *The School Review*, XLIII (February, 1935), pp. 119-131.

In an earlier article the authors reported that more than 60 percent of the Pennsylvania high schools in third class districts offer a general curriculum. It is the purpose of this report to offer data on how high school pupils use the general curriculum: Reasons why the pupils follow the general curriculum, the number of credits presented for graduation by general curriculum seniors, the scholarship of general curriculum pupils, etc.

CHILD WELFARE, GUIDANCE, AND DEVELOPMENT

BOYNTON, PAUL L., AND MCGAW, BONNIE H. "The Characteristics of Problem Children," *The Journal of Juvenile Research*, XVIII (October, 1934), pp. 215-222.

From a study of the records of three hundred and thirty pupils presented by sixty-six teachers in the fifth and sixth

grades of the Nashville, Tennessee, public schools the authors present a list of the characteristics of problem children: inattention, carelessness, lack of interest in work, will not study, disturbs class routine, whispering, etc.

BOYNTON, PAUL L., AND OTHERS. "The Emotional Stability of Teachers and Pupils," *The Journal of Juvenile Research*, XVIII (October, 1934), pp. 223-232.

From a study of some 1,095 students reported by 73 teachers in the Nashville, Tennessee, public schools the authors conclude that the study when looked at in its entirety gives very definite, clear-cut evidence that emotionally unstable teachers tend to have associated with them children who tend toward instability.

HOEFER, CAROLYN, AND HARDY, MARTHA C. "The Rôle of Health in the Child's Development," *The Elementary School Journal*, XXXV (February, 1935), 423 pp.

The author presents data on health as related to later physical growth, as related to intelligence and school achievement, as related to school behavior, as related to early developmental history and as related to certain aspects of home background. The authors conclude that the results of this investigation are thrown into clear relief in demand for changes in the physical education program.

LINCOLN, EDWARD A. "Methods and Results in the Harvard Growth Study," *The Harvard Teachers Record*, V (Cambridge, Massachusetts: Harvard University Press, February, 1935), pp. 24-33.

The author presents in this article a general survey of the methods, materials and results of the Harvard Growth Study, which is now in its twelfth year.

MEASUREMENTS, STATISTICS, AND METHODOLOGY OF SCIENTIFIC RESEARCH

MEYER, GEORGE. "An Experimental Study of the Old and New Types of Examination: II Methods of Study," *The Journal of Educational Psychology*, XXVI (January, 1935), pp. 30-40.

In a previous article the writer presented the results of a study of the effects of the examination set on the memory for certain types of sense material. In this paper the writer presents an analysis of the methods used by students in preparation for the four types of examinations: True-false, multiple choice, completion, and essay.

TYLER, RALPH W. "Evaluation: A Challenge to Progressive Education," *Educational Research Bulletin*, XIV (January 16, 1935) pp. 9-16.

The author discusses the importance of the evaluation in the eight year Progressive Education experiment and offers certain comments and suggestions for respective evaluation.

WATSON, GOODWIN, AND FORLANO, GEORGE. "Prima Facie Validity in Character Tests," *The Journal of Educational Psychology*, XXVI (January, 1935), pp. 1-16.

The authors present data on the validity of some six hundred twenty-nine items taken from thirty-one character tests and scales wherein the validity is determined by the composite judgment of one hundred fifty graduate students

in a course in the Psychology of Character.

MISCELLANEOUS

HAWTHORNE, J. W. "The Effect of Improvement in Reading Ability on Intelligence Test Scores," *The Journal of Educational Psychology*, XXVI (January, 1935), pp. 41-51.

From a study of one hundred four pupils in grades five to twelve of average intelligence and showing as much as three years retardation in reading; improvement in reading was found to bear very little relationship to improvement in intelligence test scores.

WAGNER, MAZIE EARLE, AND STRABEL, EUNICE. "Predicting Success in College Physical Sciences," *Science Education*, XIX (February, 1935), pp. 4-9.

From the study of the records of some 500 students the authors conclude that the New York State Regents Examination marks to predict college performance better on the average than the American Council Psychological Examination for high school graduates and college freshmen and the Iowa Content Test in the several High School subjects.

RICHARDSON, H. D. "The Intelligence Quotient and Secondary School Guidance," *The School Review*, XLIII (January, 1935), pp. 49-59.

From a study of the results secured from the administration of the Terman Group Test of Mental Ability to 193 pupils in the 8th and 11th grades of the Deerfield-Shields Township High School the author found that the I. Q. is relatively constant but inappropriate for guidance purposes.

VOEGELIN, CHARLES E., AND ADAMS, SYDNEY. "A Phonetic Study of Young Children's Speech," *Journal of Experimental Education*, III (December, 1934), pp. 107-116.

From a study of the speech behavior of children enrolled in the Institute of Child Welfare Nursery School at the University of California the authors present data relating to pronunciation and phonetic difficulties. Bibliography.

WELLMAN, BETH L. "Growth in Intelligence under Differing School Environment," *Journal of Experimental Education*, III (December, 1934), pp. 59-83.

From a study of the I. Q.'s of transfer children the author concludes that children in one school system may develop differently intellectually from children in other schools and that intellectual development of elementary and high school pupils may vary with the type of school experiences which they have.

DEWEY, JOSEPH C. "The Acquisition of Facts as a Measure of Reading Comprehension," *The Elementary School Journal*, XXXV (January, 1935), pp. 346-348.

From an analysis of the test results of tests of facts and inferential thinking the author concludes that pupils may recall facts without understanding them.

HANDY, UVAN, AND LENTZ, THEODORE. "Item Value and Test Reliability," *Journal of Educational Psychology*, XXV, pp. 703-708.

The authors present data indicating that items of high differentiating value yield a higher reliability than those of low differentiating value.

HEILMAN, J. D. "The Evaluation of the Average Mark of a Transfer Student," *Educational Administration and Super-*

vision, XX (December, 1934), pp. 687-693.

From a study of data secured from some thirty-five teachers colleges the author finds a considerable variation in marks. The median marks of twelve freshman classes differed almost as much as one full step of a five-point system.

MEYER, GEORGE. "An Experimental Study of the Old and the New Types of Examination: I. The Effect of the Examination Set on Memory," *Journal of Educational Psychology*, XXV (December, 1934), pp. 641-661.

After summarizing the previous investigations the author presents the results of his own investigations with 124 students in elementary psychology at the University of Michigan from which he concludes that the examination set which the individual has is of fundamental importance both in the learning of and memory for sense material. The rate of forgetting in the case of recognition tests is more rapid for learning with recognition examination set than for learning with recall examination set.

MILLER, L. W. "A Critical Analysis of the Peterson-Thurstone War Attitude Scale," *Journal of Educational Psychology*, XXV (December, 1934), pp. 662-668.

The author presents data of the Peterson-Thurstone War Attitude Scale showing that there are large variations in the numbers of subjects checking each item. The author concludes that such scales indicate a need for critical evaluation.

REITZ, WILHELM. "Forecasting Marks of New Plan Students at the University of Chicago," *The School Review*, XLIII (January, 1935), pp. 34-48.

After presenting a brief summary of previous investigations the author presents data on a study of the problem from which he concludes that under the condition of the new plan at the University of Chicago that rank in High School class is the poorest predictive measure, that average high school marks is next and the American Council Psychological Examination the best single predictor of marks on the comprehensive examinations.

BYRNS, RUTH, AND HENMON, V. A. C. "Entrance Requirements and College Success," *School and Society*, XLI, Number 1047 (January 19, 1935), pp. 101-104.

From a study of the records of 687 University of Wisconsin students the authors conclude that while foreign language and mathematics entrance requirements may be justified that they have probably been valued out of proportion to their significance for success in college.

MEAD, A. R., AND ORR, M. L. "Second Annotated List of Manuals for Observation, Participation and Student-Teaching," *Educational Administration and Supervision*, XX (December, 1934), pp. 659-670.

In the January, 1929, issue of *Educational Administration and Supervision*, the author presented a list of twenty-four manuals for use in observation, participation and student-teaching. The present article presents an annotated bibliography of some 46 additional titles.

MCNEELY, JOHN H. "Degrees Conferred by Private Institutions," *The Journal of Higher Education*, V (November, 1934), pp. 436-442.

The author presents a study of state supervision of the degree-conferring

right of privately controlled colleges and universities.

SMITH, DORA V., AND OTHERS. "Recent Research in English," *The English Journal*, XXIII (November, 1934), pp. 718-731.

This article contains a summary of investigations relating to the teaching of English, prepared by the committee on research of the National Council of Teachers of English.

TOUTON, FRANK C., AND OTHERS. "Theses Written in California Universities in Candidacy for Graduate Degrees in Education during the Academic Year 1933-1934," *California Journal of Secondary Education*, X (October, 1934), pp. 58-67.

The author presents a list of theses written in California institutions of higher learning in candidacy for graduate degrees in Education during the academic year 1933-1934.

NEW TESTS

CLARK, WILLIS W., AND LEE, J. MURRAY. *The Lee-Clark Reading Readiness Test.* Hollywood, California: Southern California School Book Depository, 1931.

The *Lee-Clark Reading Readiness Test* is designed to predict accurately the child's preparedness to learn to read. The test is constructed for use in the first grade to enable the teacher to choose the type of work that the child is capable of handling.

CLARK, MARION G. *Exercises in the Use of Historical Evidence.* New York City: Charles Scribner's Sons, 1934.

The Clark Exercises are designed for grades 3, 4, 5, and 6 and are exercises dealing with the use of historical evidence. There are two forms of the test available.

KELTY, MARY G., AND MOORE, NELLE E. *Test of Concepts in the Social Studies.* New York City: Charles Scribner's Sons, 1934.

The Kelty-Moore Tests, like the *Wesley Tests*, represent an effort to measure an understanding of terms in the social studies course. These tests are constructed in the main for the intermediate grades of the elementary schools. They may be used in the junior high school.

The Terman Record Form. New York City: Houghton Mifflin Company, 1934.

A record form designed for use with the *Stanford Revision* of the *Binet-Simon Tests*, copyright, 1916, 1922, and 1934 by Houghton Mifflin Company.

WESLEY, EDGAR BRUCE. *Wesley Tests in Social and Political Terms.* New York City: Charles Scribner's Sons, 1932.

The *Wesley Tests* are examinations in social and political terms and were constructed for the Commission on the Social Studies in the schools. The tests are designed to measure achievement in the usage and understanding of the social and political terms used in the social studies of the present day schools. They are designed for the senior high school level.

RESEARCH NEWS and COMMUNICATIONS

Address all research news and communications to Carter V. Good, Teachers College, University of Cincinnati, Cincinnati, Ohio.

Yearbook for 1935. — The Thirty-Fourth Yearbook of the National Society for the Study of Education, entitled "Educational Diagnosis," has appeared in a single volume. It was prepared by a committee of the Society composed of L. J. Brueckner, University of Minnesota (Chairman); G. T. Buswell, University of Chicago; W. C. Olson, University of Michigan; P. T. Rankin, Detroit Public Schools; J. A. Stenquist, Baltimore Public Schools; L. E. Travis, State University of Iowa; and R. W. Tyler, Ohio State University, assisted by several other members of the Society and a few specialists from without the Society.

The scope of this important volume is indicated by citing certain of the topics treated: factors associated with learning difficulty (physical, intellectual, pedagogical, social, environmental), principles and technics of diagnosis, diagnosis and remedial treatment in various fields and objectives of instruction (reading, English, arithmetic, social studies, science, health, speech, vocational interests, music, art, recreation, creativeness), and administration of a program of diagnosis and remedial instruction. It is obvious that these topics are matters of direct concern to superintendents, supervisors, principals, and classroom teachers. The presentation, likewise, will be of assistance to college teachers of education and to specialists in educational research.

The following yearbooks of the Society have been initiated under the chairmanship of the person indicated: "International Relations and the School Program" (J. T. Shotwell), "Organization of Higher Education in the United States" (L. D. Coffman), "Education as a Science" (H. Rugg), and "Music in the Public Schools," (W. L. Uhl). It is impossible at this date to state the order in which these yearbooks will appear in print.

Carnegie Corporation.—The 1934 report of the president (F. P. Keppel) and treasurer of the Carnegie Corporation of New York shows that, of almost exactly $5,000,000 voted from the income of two funds during 1933-34, some $3,688,320 was dependent on policy decisions arrived at in earlier years. The largest items are as follows:

For the permanent funds of the American Library Association	$500,000
To the American Law Institute for the Re-statement of the Common Law	538,000
For support of the Institute of International Education, for a second five-year period	300,000
For the continuation of the Australian Council for Educational Research and for the establishment of a similar institution in New Zealand	250,000
For support of the American Association for Adult Education and the National Occupational Conference	238,000

Following British Empire surveys as to workers' education and museum development.....$136,250
Following London conferences of 1933 on scientific research in the British Empire..... 166,500
Continuation of other programs, involving relatively smaller grants 759,820
Continuation of coöperative programs with other Carnegie organizations 799,750

New undertakings include an endowment grant of $100,000 to the Museum of Modern Art, $200,000 divided among the universities of Western Canada, and grants from the British Dominions and Colonies Fund, aggregating $205,000, to finance a general survey of Southern and Equatorial Africa, to contribute toward social research in the Union of South Africa, and to support a three-year program in the interest of over-seas students at the Institute of Education, University of London.

The Corporation has done important work in awarding fellowships, travel allowances, grants-in-aid, and exchange visits. Two general studies are under consideration, one on support for scholarly publications, and one on grants to operating organizations for general expenses as contrasted to grants for specific programs.

The Corporation has made possible the work of the Commission on the Social Studies and the Pennsylvania study of articulation between secondary schools and colleges, as well as the earlier studies of modern foreign languages, engineering education, forestry, architecture, and the recently completed survey of dental education.

International Education.—The fifteenth annual report (October, 1934) of the Director (Stephen Duggan) of the Institute of International Education includes a pointed statement of the detrimental effect of militarism on culture in Europe. The Institute serves in a variety of capacities in promoting cultural activities internationally, which are described in the annual report: representative in the United States of summer sessions of foreign universities, an emergency committee in aid of displaced German scholars, exchange lectureships, preparation of data concerning American professors who are to go abroad on sabbatical leave, and exchange fellowships and scholarships. The Carnegie Corporation has renewed an annual grant of $60,000 for a period of five years for the support of the Institute. The $24,000 a year granted by the Rockefeller Foundation still has four years to run. The Woodrow Wilson Foundation added $4,000 during the year to the resources of the Institute.

Rockefeller Foundation, 1933.—During 1933 the Foundation appropriated $9,890,806 for various philanthropic projects in the field of: public health, $3,286,063; medical sciences, $1,173,853; natural sciences, $807,250; social sciences, $1,636,000; humanities, $847,500; and emergency grants, $589,000.

In the words of Max Mason, president of the Foundation, the emphasis of the organization has been as follows:

"While the Foundation has continued during the year its activity in furthering the advance of knowledge in the fields of public health, medical science, natural science, social science, and the humanities, its work in these fields has been vitally changed in scope through selection, for intensive work, of those subfields which contribute more directly to the general problem of human behavior,

with the aim of control through understanding.

"In the fields of medical and natural science the major emphasis has been placed on the problem of mental health and on the development of those sciences whose advance is imperatively demanded to form the substantial scientific basis for the rapidly evolving modern science of man. In the social sciences the problems of economic structure and process, international relations, and community organization and planning have been considered of principal importance. The emphasis in the humanities has been on the encouragement of international cultural understanding and the preservation and interpretation of American culture. In the International Health Division, there has been increased attention to studies of disease in its environment, with closely related laboratory investigations.

"Early in 1933 a fund was set aside for projects having to do with the present national and world situation. The projects which have been supported fall under two general headings: studies providing for concurrent appraisal and studies furnishing a factual basis for improved planning of important sections of the recovery, relief, and reconstruction program; and undertakings involving the collaboration of private agencies in the prosecution of the emergency programs of the federal, state, and local governments.

Another emergency requiring special funds was the displacement of eminent scholars in Europe for political reasons and the consequent interruption of important scientific work. Dealing directly with universities in Europe and America which wished to add to their staffs, the Foundation has contributed toward their salaries. The situation was an emergency; the solution of the problem is as yet on a temporary basis."

Curriculum Bibliography, Courses of Study, and Textbooks.—The *Society for Curriculum Study News Bulletin* for October 24, 1934, includes the sixth (March, 1933, to March, 1934) annual bibliography on curriculum making prepared for the Society. The work was done by Henry Harap of Western Reserve and by Edgar Dale and Vivian Weedon of Ohio State. The Bibliography numbers 246 items. The literature in this field is characterized as follows for the year represented:

The number of general articles and monographs on curriculum making continues to increase, a tendency which was also noted a year ago. The output of investigations of curriculum objectives continues to decline. There has been a considerable increase in papers pertaining to the nature of the unit of work. Interest in activity curricula appears to be sustained. There has been a marked increase in the number of entries in the fields of higher education, teacher training, and social studies.

The January 15, 1935, number of the *Bulletin* contains a list of the outstanding courses of study of 1934 (a total of 142), compiled by H. B. Bruner of Teachers College, Columbia University, and a list of the textbooks (a total of 460) of 1934 (elementary, junior-senior high school, and junior college), prepared by M. E. Herriott of the Los Angeles schools and his committee.

Distribution of Graduate Theses, 1932-33.—Office of Education Bulletin No. 7, 1934, lists graduate theses and faculty research in Education for the twelve months ending August 31, 1933. The distribu-

tion of theses is as reported in Table I, indicating that the fields of greatest interest are the same as for the school year, 1931-32. The library of the Office of Education has more than 1,300 of these graduate theses available for loan. This is the seventh compilation of research issued by the Office.

TABLE I

NUMBER OF THESES IN VARIOUS FIELDS OF EDUCATION, 1932-33

Subject	Doctors' Theses	Masters' Theses
1. Education — history and biography	19	105
2. Current educational conditions — United States	3	38
3. Current educational conditions — foreign and international	12	36
4. Educational theory, special methods	21	142
5. Psychology, educational and child study	11	61
6. Testing and research	14	130
7. Curriculum studies, including subjects of the curriculum	111	1,285
8. Elementary education, including preschool	8	59
9. Secondary education and junior college	22	161
10. Higher education	22	121
11. Adult education	1	24
12. Teacher training and teachers, status	34	137
13. Administration of schools and school management	56	648
14. School buildings, equipment, and janitors	5	33
15. Sociology, educational	3	29
16. Racial groups, education	9	61
17. Exceptional groups	11	89
18. Libraries	8	70
Total	370	3,229

Training for the Doctorate. — In the 1934 *Report of the Dean of Teachers College*, Columbia University, W. F. Russell describes a challenging plan for training advanced students in this country, based on observation of the program employed in the University of Paris, with a new degree of Doctor of Education added to the existing Ph. D. programs:

"Thus from this famous school, the mother of the educational élite of France, we learn an interesting lesson. If you can assure the ablest students of a permanent post for life, you can attract thousands to turn their hopes in your direction. If, in addition, you can eliminate the economic barrier, by providing all expenses over a period of years, you can be assured that the ablest will take the examination. Even if your examination system is imperfect, and even if you cannot be sure that you judge either fairly or accurately, none the less you will have a first-rate body of students. Then, if you provide a place for study, a library, and laboratories, and if there is a rich offering of lectures and classes in other educational institutions near-by to which you can send your students, and if, in addition, you have an adequate system of supervision and periodic examinations, then you need have no faculty of your own.

"I should like to see something like this at Teachers College. With courses arrayed under five divisions in the Faculty of Teachers College, could we not superimpose a small Advanced School of Education for the development of an educational élite? To what degrees could the system of the Ecole Normale Supérieure be incorporated in our institution under American conditions?

"In the first place we cannot assure any person a permanent post in civil service. Certainly this has been, and is, contrary to American educational practices; but there is no doubt that if we were to have a rigid system of selection, all the students on whom we should confer the Doctor's degree would be assured of posts of dignity and importance. Of course under present conditions we could not offer free maintenance, but at least we should consider whether or not it might not be worth while to try to raise three millions of dollars, the income from which could be used for this purpose. The government follows this plan at West Point and Annapolis to provide the future officers of the Army and Navy. The general public subsidizes most students at theological seminaries and medical schools. For the national security we must have properly trained leaders in warfare and religion and medicine. Have we any lesser need of leaders in education?

"Of course we could not have general competitive examinations under government regulation. But graduates of Teachers College are in key positions everywhere, both in the United States and in other countries; and each year in regular classes, in the Summer Session, and in extramural and special classes we come into close relationship with some 20,000 different students. Is there not among this number the raw material of an educational élite?

"I am certain that the Faculty of Teachers College, so successful in devising examinations for other agencies, could devise some form of test by means of which we could select the best students for the new school. Even without great financial encouragement, many of the best, when selected, would be able to pursue their advanced studies.

"Furthermore, we could follow the French plan in the new organization of Teachers College. We could establish an Advanced School of Education on the model of the Ecole Normale Supérieure, with the rest of Teachers College, the University, Union Theological Seminary, the Juillard School of Music, and other Institutions in the same relation to it as the University of Paris is to the Ecole Normale. Under this system we should need a place for the students to gather and to work, a library, special laboratories, and no faculty at all which would belong to it exclusively. Professors in Teachers College could be assigned to it for full time or part time. Young graduates of highest promise could be retained on three-year terms to work with these students. The students could spend their time anywhere in the University, and, if deemed wise, at the Sorbonne, University of London, Harvard, Yale, Chicago, the Brookings Institution, and similar educational centers.

"Sometime this school should have a great building, facing the College on 122nd Street. There should be a spacious social hall with a wood fire burning and plenty of books. There should be a large auditorium for general lectures. There should be smaller conference rooms, numerous offices, a great study hall, a library, and quarters for statistical, curriculum, and guidance laboratories. There

should be living quarters for students and staff. Here would be a modern Solomon's house, accepting only the best, graduating only the most able, attacking on all fronts the problems of education. It would be an honor to be a member — a privilege highly to be appreciated."

Washington Experiment in Character Education.—The *Educational Record* for October, 1934, contains an interesting description, written by W. W. Charters, of a significant attempt in the District of Columbia to correct anti-social tendencies in children.

"In the United States today the average age of the prison population is 23 years. The largest age group is found at 19 and the next largest group at 18. Most criminal careers begin in childhood.

"The foregoing facts are three of the most significant findings established by the hearings conducted by the United States Senate Sub-Committee on Racketeering and Crime last year. These facts led the Chairman of that sub-committee, Senator Royal S. Copeland of New York, to institute a campaign for strengthening the public schools in their treatment of anti-social tendencies in children. His challenge to the schools was delivered to the Department of Superintendence of the National Education Association at its meeting last February (1934) in Cleveland, and to the American Council on Education at its meeting last May in Washington. Both of these challenges have been given wide publicity.

"Having spent most of his life in medicine and public health, Senator Copeland fully understands the enormous advance that has been made in these fields by clinical methods of work. Hence he pointedly asks, why not approach the attainment of mental, moral and social health in much the same way as we approach that of physical health? The doctor deals with patients as individuals. Why should not the teacher deal with the children as individuals?

"This shift of emphasis from obedience to the mass prescriptions of a required curriculum to prescriptions adapted to the individual needs as shown by appropriate diagnosis and behavior records, Senator Copeland terms the clinical approach to education. His specific proposal to the schools is that they use every resource of research and experimentation to develop this clinical method of treatment as rapidly as possible.

"With the enthusiastic approval of the Board of Education of the District of Columbia the experiment proposed by Senator Copeland has been undertaken in the public schools of Washington under the direction of Superintendent Frank W. Ballou. In June, 1934, the Seventy-Third Congress added approximately $70,000 to the ten million dollar budget for the public schools of the District of Columbia at the request of Senator Copeland and the Board of Education for this purpose. Washington was selected for this study, partly because it was a convenient location, partly because Superintendent Ballou and the Board of Education are enthusiastic about the project, and partly because the teaching corps of the city has been working upon the problem for the last five years. The purpose of the appropriation is to provide for adequate experimentation to discover how the desired results can be secured at the lowest cost.

"In setting up the program, it was decided to use a spearhead attack, or, in other words, to work in a small area most favorable to success, make it as effective as possible, learn from this experience and extend into other areas as soon as practicable. In following this

plan five white schools and five colored schools were selected on the elementary, junior high, senior high, night high, and vocational school levels. The fourth grade in the elementary school was brought into the project, the seventh grade in the junior high school, and the tenth grade in the senior high school. In 1935 the fifth, eighth, and eleventh grades will be added, and in 1936, the sixth, ninth, and twelfth grades will also be included. In 1934-35, approximately 175 teachers and 5,500 pupils are directly included in the project."

Reading Disabilities.—The National Conference on Research in Elementary School English has sponsored the preparation of a research bulletin on "Reading Disabilities" which was presented at the Atlantic City meeting of the Department of Superintendence. Abstracts of investigations pertinent to the problem are included. Among the topics considered are: administrative of a reading clinic, analysis procedure, auditory characteristics of good and poor readers, auditory readiness, auditory span, binocular aspects, concepts in reading ability, corrective procedures, dominance, emotional factors, eye movements, handedness, hygiene of reading, individual differences, information tests, kinasthetic factors, kindergarten requirements, laterality, leading, length of line, lighting research, maturation in relation to reading achievement, memory span, mental hygiene aspects, mirror reading, pedagogical procedures, perception habits, personal factors, photographic measures, physiological aspects, reading aptitudes, reading methods, reading readiness, size of type, sentence structure and comprehension, sight saving, typography, visual readiness, visual span, and vocabulary burden.

Phi Delta Kappan.—The journal, *Phi Delta Kappan*, appeared with a new format and a somewhat changed policy with the November, 1934, issue. This publication will remain the official organ of the professional educational fraternity, Phi Delta Kappa, with a broader editorial policy, nine issues a year, and with the same editor, Paul M. Cook. Added features or sections are "Editorial Comment by Contributing Editors," "Keeping Abreast of the Times," "Improving Practice through Research," and an "Open Forum."

Status of High School Graduates in Minnesota.—In 1933 there was a total of 21,163 high school graduates in Minnesota. One year later, data were reported for 18,227 of these graduates in terms of whether they were: (1) continuing their education, (2) engaged in an occupation, or (3) unemployed. In terms of percentages the facts were as follows: attending some institution of learning, 35.67; employed in some occupation, 41.16; unemployed, 14.66; married, 1.93; deceased, .14; no information reported, 6.44. The study was directed by T. J. Berning of the State Department of Education.

National Occupational Conference.—The National Occupational Conference is a non-commercial organization which, in addition to other activities including research and publication, has taken over and is continuing the field service formerly conducted by the National Vocational Guidance Association. It is supported by a grant from the Carnegie Corporation. The Conference provides for educational institutions, libraries, and other interested organizations a consulting service regarding the theory and practice of vocational guidance, and the re-

sults of research in occupational adjustment. Upon request, a staff officer will visit local institutions for consultation regarding the organization of work designed to contribute to the better occupational adjustment of more than one person. The Conference will contribute the services of staff officers for such trips, but will expect local institutions to defray the necessary expenses. No charge is made for any assistance which can be given by mail. The work of the Conference does not include counseling with individuals regarding their personal occupational problems. The organization is located at 522 Fifth Avenue, New York City.

New Journal.—A group of workers in the social sciences is planning tentatively for the establishment of a *Journal of Social Philosophy*, a quarterly to be devoted to a philosophic synthesis of the social sciences. It is hoped in part to fill the gap left by the discontinuance of *Social Science Abstracts*. The secretary of the proposed editorial board is M. J. Aronson of the College of the City of New York.

Research in Secondary Education.—A brief statement of the history and work of the National Committee on Research in Secondary Education, organized in 1925, recently has been distributed in mimeographed form. The present officers are E. J. Ashbaugh, Miami University, chairman; W. A. Wetzel, Trenton, N. J., public schools, vice-chairman; and C. A. Jessen, U. S. Office of Education, secretary. The work of the Committee has been evaluated by G. F. Zook, Director of the American Council on Education and former U. S. Commissioner of Education, as follows:

"The National Committee on Research in Secondary Education has in my judgment performed an important service to education, especially in its activities looking toward stimulation and coördination of research. I am one of those who believe that the National Committee on Research can exert a more significant influence on education through the contacts of its members than through any production schedule in original research which it might set for itself as an organization. Because its members are busy men in important positions they cannot in general give to the Committee's work the large blocks of time needed in the production of research studies. They can give, and have given, without remuneration, time and effort needed in stimulating others to coördinated attack on problems in secondary education. The relationships of the Committee to the Tercentenary Celebration of the founding of secondary education in the United States, to the production of bibliographies of research studies in education, and to the National Survey of Secondary Education are good illustrations of what I have in mind. The Committee's assistance in starting and following up important undertakings of this sort may not be recognized in volumes or other concrete evidence bearing the Committee's name, but the contributions are nevertheless real and essential.

"I hope that the committee may be willing in the future to take upon itself activities involving long-scale planning and the forming of policies in the field of secondary education. In this connection, I commend the special interest which the Committee has recently evidenced in the problems surrounding the education of youth. When we reflect that more than half of the members of our population between the ages of 14 and 20 are not in school and that relatively few of them are employed, it becomes apparent

that success can be attained only through the marshalling of all educational agencies both within and outside the school for a concerted attack on the youth problem and through the supplementing of the work of these agencies as necessary. In my opinion this problem offers a challenge to the best thought of our best thinkers."

Coöperative Test Service and Educational Records Bureau.—The purpose of the Coöperative Test Service, under the terms of the subvention from the General Education Board, is to construct ten or more comparable forms of examinations in the fundamental subject-matter fields of junior college and senior high school levels, and to make them available to the colleges and schools, one form each year, at the lowest possible cost. Colleges and schools may inaugurate the use of these tests with the assurance that a new comparable form of each will be made available each year about April 15. Some of the advantages of such a series of comparable examinations are that:

1. Individual growth in defined types of achievement can be measured year after year, thus making feasible types of educational guidance and of educational research which would be difficult or impossible with unrelated examinations or with "standardized" tests which exist in two or at the most three comparable forms.

2. Each college, school, or school system that uses these tests may set up its own standards for admission, placement, promotion, certification, and graduation, and maintain those standards uniformly from year to year, and at the same time maintain transfer and advanced standing relations with other institutions on the basis of comparable achievement test measurements.

3. The results of subjective and other local examinations can be made closely comparable from year to year by using the results of the Coöperative tests as a common denominator, thus taking advantage of the best features and minimizing the weaknesses of both types of examinations.

The Educational Records Bureau, a non-profit making service and research agency for schools and colleges, was organized in the fall of 1927 to assist member schools in securing reliable and comparable measurements of the native capacities and academic achievements of their pupils. Such measurements, based on comprehensive objective tests over a series of years, aid materially in a continuing study of pupils, and are recognized as an indispensable supplement to other types of information currently used in administering transfers, promotions, admissions to secondary schools and to colleges, and in other aspects of constructive educational guidance.

The coöperative nature and the comprehensive scope of the annual achievement and academic aptitude testing enterprises sponsored by the Bureau help participating schools to obtain the maximum guidance and administrative values from coördinated continuous programs. Service is rendered more effectively by the Bureau's close association with the Coöperative Test Service and the Committee on Educational Testing of the American Council on Education. Ben D. Wood of Columbia University is director of both services.

Australian Council for Educational Research.—Excluding renewals and extensions of existing grants the Council, since its establishment, has undertaken to sup-

port 95 distinct inquiries. Of these, 44 have been completed, 34 are still in progress, 11 have not yet begun, and 6 have been abandoned. The total sum expended by the Council in subsidizing investigations, apart from inquiries it has itself originated, amounts to 10,250 pounds.

To date the Council has published 37 separate reports.* Of these, 16 were manuscripts submitted by persons who had not received a grant from the Council for carrying out the work concerned. It is thus obvious that the facilities provided by the Council have served a useful purpose not only in stimulating inquiries but making it possible for authors to publish work of educational importance.

Chico, California, Projects.—Studies at Chico State Teachers College include: population trends in the counties served by the Chico State Teachers College (G. A. West); upper-grade reading in the college demonstration school (F. F. Caldwell); the respective merits of teacher-training programs carried on in schools off the campus and in the college demonstration school (L. D. Bernard); and certain studies in connection with a personality inventory, recently published by the Stanford University Press, which attempts to measure home, social, health, and emotional adjustments (H. M. Bell). The relationship of the four sections of the Inventory to intelligence, scholarship, success in teaching, delinquency, and social and economic status are being determined.

Research Projects at Evanston, Illinois.—The National College of Education, in response to a need for data other than the academic standing of college students, has organized a detailed list of 170 positive traits subdivided under the following topics:

1. Physical Traits
   - General
   - Personal
2. Social Traits
   - Coöperation
   - Responsibility
   - Initiative-Originality
   - Leadership
   - Self-Control
     - Voice
     - Emotions
   - Disposition-Personal Traits
3. Intellectual-Educational
   - Class Assignments
   - Class Discussions
     - Participant
     - Listener
   - Work Periods
   - Personal Interviews
   - Committee Work

For the past few years members of the Department of Guidance have been assembling tests which will differentiate the skill of one hand or foot over its mate.

Work has been done on additional tests for eye preference, eliminating the manoptoscope, since it seemed unreliable when used with children. Some interesting clues have been discovered by having a child read for three minutes with first one eye alone, then the other, and finally the two eyes simultaneously. Apparently, some children have been discovered who read much more contentedly and more accurately with one eye alone than with the two eyes simultaneously. Further tests are now being devised for eye and ear preference, and eye and ear skill, in a coöperative study with W. S. Gray at the University of Chicago.

In the Children's School, pupils are placed according to social and mental maturity, and in the case of those few children who have not progressed nor-

mally in their academic subjects, individual instruction is given, based upon clinical findings. This instruction is given by a trained assistant or by upper classmen taking courses in clinical procedures, and is supervised by the instructor and assistants in these classes. In computing the achievement quotients for nearly one hundred children from the second grade through the seventh, the achievement quotients and percentages of pupils represented are respectively: 100 or over, 51.6; 95-99, 22.6; 90-94, 12.9; 80-89, 10.7; below 80, 2.2. It should be stated that this is a school of superior children mentally and socially, organized in small groups of not more than twenty-five pupils.

Louise Farwell is Director of Research and Guidance.

Minneapolis Curriculum Study.—Curriculum Bulletin No. 90 of the Minneapolis schools outlines the proposed program of the Instructional Division for the current school year. For each of the school subjects the extent of progress made is reported briefly and problems for further study are stated. Prudence Cutright is Director of Curriculum.

Follow-up Study in Missouri.—J. D. Hull, principal of the senior school at Springfield, Missouri, is directing a follow-up study of the pupils who graduated from this school during the years 1929, 1931, and 1933. The purpose of the investigation is to gather information which will be helpful in the guidance program of the school. Five students of the Southwest Missouri State Teachers College at Springfield have been assigned by the M. E. R. A. to carry on the work. These students will interview as many of the fifteen hundred graduates as possible and question them concerning their experience and interests.

Investigations at Illinois State Normal University.—Projects recently completed at this institution include: subject combinations in Illinois high schools, participation in the development of an intelligence test for the Teachers' College Personnel Association, a time-controlled exposure apparatus invented and patented by C. F. Malmberg, types of administrative procedure in teachers college, and the nature of absences in the local institution.

Philippine Public School System.—(1) Relative Effectiveness of the Two Single-Session and the Whole-Day Plans of Instruction—Under the two single-session or two half-day program of instruction each pupil attends school in accordance with a shortened program during a half-day session only, either during the morning or during the afternoon. The teacher, however, teaches during the entire day—one class in the morning, and a different class in the afternoon.

Under the current whole-day program each pupil attends school both morning and afternoon. The work is broken up into the morning and the afternoon sessions. As in the two single-session program, the teacher teaches both morning and afternoon, but has the same class.

To show the relative effectiveness of the plans, two experiments were conducted, one serving as a check on the other. The results in both experiments showed: (a) that the pupils under the current whole-day program learned more than the pupils under the two single-session program, as revealed by the test scores received in both the Bureau of Education standard tests and in the locally prepared tests; (b) that a high percentage of promotion among the pupils was made under the current whole-day program; and (c) that the teachers under the two single-session plan showed signs of an impaired health, vigor, and vitality.

2. Relationship between the Teacher's Instructional Skill and His Personal and Social Qualities.—Instructional skill includes such items as daily preparation, technique, class control and management, use of English, evidence of reading and study, and interest in pupil activities. Under personal and social qualities come physical condition, industry, initiative, coöperation and loyalty, punctuality, character, and standing with pupils, fellow teachers, and community. The sources for this study were the data contained in the Bureau of Education Form 8-A, Efficiency Report for Classroom Teacher, for the school year 1933-34 submitted by the supervising teachers in the province of Samar. This study included more than one-half of the teachers of the province. The results showed a high correlation between instructional skill and personal and social qualities of teachers, the coefficient of correlation being .86 and the P. E. ± .01.

3. School Subjects in Which the Greatest Number of Pupils Failed.—In order to obtain the data necessary for this study the Bureau of Education Form XVIII-B, Report on Promotions for Grades IV to VII, for the school years 1928-1933 inclusive, and for all municipalities of the province of Oriental Negroes was examined. The results revealed that the greatest number of failures in Grades IV to VII was in arithmetic. Geography ranked second in Grades IV to VI, and Philippine history and government in Grade VII. Spelling, phonics, and conversational English showed the lowest mortality.

4. A Study of Verbs in Grade V Written Composition.—Of the 25,107 verbs scored, 8,483, or 35.22 percent, were erroneous. Nearly three-fifths of the number of errors were the past tense. Tables were made showing the distribution of the 8,483 errors and the percentage of errors, the distribution and percentage of errors with the various verb forms, with the different tenses, with infinitives and participles, with the omission of the direct objects, and with the use of *bring* and *take*.

5. A Study of Spelling.—The 2,774 words most commonly used were arranged in separate lists according to the size of the individual frequency of each word. Of these 2,774 words, 1,504 which are not in the Bureau's courses of study were marked. A list of 1,194 words found in the courses of study, but which had a frequency of less than three in the running words met in the study, was made.

6. Grade IV Spelling Difficulties and Corrective Exercises.—A list of words most frequently misspelled (together with the types of errors made in each case of misspelling), words written as one, and words misused interchangeably was made. Corrective exercises, based on the errors found, were prepared.

7. Errors in English and Mechanics of Writing Made by Grade IV Pupils.—This study aimed to determine what errors predominate in the use of English. It was found that certain points in connection with mechanics in writing were overlooked, that sentence form was often incorrect, and that frequently the wrong word form was used.

8. Errors in English and Mechanics of Writing Made by Grade III Pupils.—Two thousand one hundred and four seat-work papers of Grade III pupils from 11 school divisions were used for this study. The most outstanding errors found in English and in mechanics, and their corresponding percentages of error are as follows:

Incorrect form of third person, singular number, present or past tense	13.6

Failure to use the period at the end of a statement	10.83
Use of present tense for past tense	9.02
Failure to use plural nouns when required	8.7
Capitalizing words other than the first word in a sentence or other than proper names or adjectives	7.00

9. The Personal Pronoun in Grade IV Written Work.—From the examination of 1,100 written compositions and seatwork exercises of Grade IV pupils, it was found that the personal pronouns were used extensively and for the most part correctly, the percentage of error in their use being only 9 percent. The most prevalent errors were the interchanging of *he* and *she*, *he* and *it*, and *she* and *it*.

10. The Teaching Load of Secondary Teachers.—The instructional load of teachers varied in the different curricula. This variation lies in preparation, routine, student relations, extra-curricular activities, and work of a special nature. There is a decided relationship between the number of different preparations and the teaching load. The median in the various curricula is approximately nine hours per day

11. Study of Specific Teaching Needs and of Teachers' Special Qualifications.—To carry out the general policy of giving the teachers major assignments in the subjects in which they majored in college and minor assignments in those in which they minored, the General Office analyzed secondary-school programs and urged that new appointees be selected in strict adherence to this policy. In many instances teachers already in the service were also shifted so that, as far as local conditions permitted, each teacher was given a major assignment in his major subject. No teacher was assigned to teach a subject for which he had no adequate preparation or in which he had no plans to pursue advanced studies.

In order further to insure proper placement of teacher applicants, the Bureau of Education also considered their scholastic records, qualifications to handle extra-curricular activities, command of spoken English, and teaching personality, through personal interviews. In several instances the scores of the applicants on tests administered in the General Office also were considered.

12. A Study to Determine the Apparatus Essential for Science Classes in Secondary Schools.—This study revealed the apparatus absolutely essential to efficient instruction in the different courses in secondary science. A list of essential equipment and supplies was made separately for general science, biology, and physics. The results of this study should help school administrators in placing orders for laboratory equipment in such a way as to insure maximum returns for the amounts expended.

Manuel L. Carreon is chief of the Measurement and Research Department.

Purdue Studies in Higher Education.—A recent bulletin, *A Ten-Year Review of Studies in Higher Education*, prepared by H. H. Remmers, summarizes the results secured in this field at Purdue University. The bibliography includes 23 Studies in Higher Education, 5 Educational Reference Circulars, and 34 articles published in journals. President E. C. Elliott comments as follows in the Foreword:

"For nearly a decade the Division of Educational Reference has been an integral unit of the operative mechanism of the University. From its inception the underlying purpose of the Division has been to subject selected procedures, policies and problems of higher education to scientific analysis and constructive valuation. Some time ago it appeared that the time had come for the Division

to take an inventory of its efforts and accomplishments as a first step for determining the worth of the Division for the improvement of the essential educational activities of the University. The Director of the Division has presented the results of such an inventory in the following pages.

"The primary purpose of this review of the work of the Division is to assemble and to make readily available to the members of the University staff the results of the more important studies made during recent years. At the same time it is hoped that the experience of Purdue University, in attempting to deal objectively and scientifically with its own educational methods and mechanisms, will be of service to other higher institutions that are aware of the enigma of educational efficiency."

Educational References.—The U. S. Office of Education is publishing a series of booklets of approximately ten pages each, listing "Good References" on a variety of educational problems, which may be secured without charge, upon application to the Commissioner of Education at Washington.

Wisconsin Radio Service.—In response to a frequently expressed need, WHA (Madison) offers a weekly program for Wisconsin teachers in service. Since Wednesday, October 3, members of the department of education, University of Wisconsin, have offered a series of discussions of the everyday problems of classroom instruction at 4:00-4:30 P.M.

It is hoped in this manner to make it possible for teachers in their own schools, without loss of time or the expense of travel, to hold weekly institutes upon the newest developments in the field of education. These discussions attempt to clarify issues, summarize relevant research, state guiding principles, suggest implications of such principles for the improvement of practice, and open up the possibilities of further study and research. A. S. Barr and M. H. Willing of the University of Wisconsin are in charge of these programs.

Follow-up Study of High School Pupils.—W. M. Proctor of Stanford University is completing a thirteen year follow-up of some sixteen to eighteen hundred high school pupils to whom the Army Alpha intelligence test was given during 1917-18. He has made two check-ups on this group previously. The fourth year follow-up was included in the monograph on the "Use of Intelligence Tests in the Guidance of High School Pupils," published in 1923. The second check-up was made in 1927. The thirteenth year follow-up was made in the year 1930-31. These data are now being tabulated and charts are being constructed. There will be findings with regard to the relationship between scores made on the intelligence test in 1917-18 and the occupations in which the persons were found in 1930-31. There will be studies of the relation between the fathers' occupations and the intelligence of the sons and daughters, as well as between the rank of the sons' and daughters' occupations and the rank of the fathers' occupations. It will be possible to make some deductions concerning the effect of social and economic status upon the success of given students in secondary schools and colleges, and also the level of occupational success achieved at the time the final check-up was made.

It is hoped to have this material ready for publication some time in April.